Collaborative Research: SI2-SSI: PI David Tarboton, Lead Institution: Utah State University

An interactive software infrastructure for sustaining collaborative community innovation in the hydrologic sciences

Major gaps exist in our basic understanding of water availability, quality and dynamics. Comprehensive understanding of the interactions between natural and engineered aspects of the water cycle requires integration of multi-source and potentially large-scale data and modeling, where cyberinfrastructure can play a crucial role in enabling research progress. In addressing the question, for example, of how water supply and quality will change in response to climate and land use change, a hydrologist needs to obtain information on land use, streamflow, terrain, water quality, soils, meteorology and climate predictions from multiple agencies or other universities and effectively integrate this information with field data to make predictions and test hypotheses. At the present time, most researchers perform this integration manually. This can be a daunting process and information management challenge that could be greatly simplified through cyberinfrastructure tools for collaboration and shared access to data and models. **The goals of this project are to: 1) develop sustainable cyberinfrastructure for better access to data and models; 2) to enable the hydrologic and other communities to collaborate and combine data and models from multiple sources, and to thus transform the way hydrologic knowledge is created and applied; and 3) provide new opportunities for information integration at a scale that was previously inconceivable to hydrology researchers.** The field of hydrology is diverse and requires integration of knowledge from across the geosciences, ecological sciences, biological sciences, social sciences and engineering, and so represents a fertile test case for pursuing these goals.

The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI), with over 125 member Universities has, with NSF support, developed a Hydrologic Information System (HIS) to provide and enhance access to hydrologic data. HIS includes a services-oriented architecture for publishing hydrologic data that has been deployed by multiple universities and government agencies. Achieving the goals stated above requires extending HIS through a sustainable software engineering strategy that enables sharing of both models and data as objects that can be iteratively improved in a collaboration environment and can support education of a new generation of hydrologists versed in using data intensive methods to generate knowledge. Publication services also need to be enhanced to enable curation and archiving of data and models as scientific works. In summary, expansion of HIS is needed in three key areas: *interactive data access*, *model and tool sharing*, and *education*.

Intellectual Merit. We propose to develop an online, collaborative environment for the sharing of hydrologic data and models called HydroShare. HydroShare will enable scientists to easily discover and access data and models, retrieve them to their desktop or perform analyses seamlessly in a high performance computing environment, thereby enhancing research, education and application of hydrologic knowledge. HydroShare will be a unique, multidisciplinary integration of cyberinfrastructure, community processes and software engineering that will combine existing NSF-supported cyberinfrastructure (CUAHSI HIS, iRODS, and HUBzero) to create transformative cyberinfrastructure for collaboration, data sharing and enhanced modeling. The engineering process used for the open source, user-driven development of HydroShare and the engagement of commercial and publishing partners will lead to its sustainability beyond the proposal funding.

Broader Impacts. The software infrastructure developed will provide tools to better answer water related research questions with impacts related to climate change, floods, droughts, biofuels, etc. and will improve decision making on critical societal issues affecting water. CUAHSI HIS already has a broad impact with a rapidly growing user base and has been identified in some NSF program solicitations as a solution for data archiving requirements. HIS is actively supported by CUAHSI through a full time User Support Specialist, periodic user training workshops, and promotion at national meetings such as American Geophysical Union (AGU). The CUAHSI HIS team - including new team members added with this proposal - is well versed in the technological options for building HydroShare, but more importantly – through the auspices of CUAHSI as a community consortium – have demonstrated the ability to engage and draw upon community input providing the user perspective crucial to ensure adoption and success. HydroShare will also have broad impact by serving as a blueprint for combination of software tools and collaboration applicable to other geo-informatic disciplines.

Project Description

1. Introduction

Water is essential to human life, but advances in hydrologic understanding have been hampered by inadequate cyberinfrastructure (CI) for data access, integration, simulation and education. Motivated by the need for better CI, the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) developed its Hydrologic Information System (HIS) project, with NSF support, to enhance access to hydrologic data. The problem addressed by HIS is that the data required to understand terrestrial water are spread over many incompatible data servers at federal agencies, universities, and other organizations. Hence, simply obtaining the necessary data in a usable format is a daunting challenge that slows the pace of hydrologic research. The CUAHSI HIS has begun to transform the way data is accessed and shared by the hydrology community by creating standards, server and desktop software, and a comprehensive data catalog supporting data discovery. We believe that even greater advances are within reach by creating new functionality for collaboration and interactive data and model sharing by groups of researchers.

HIS has proven effective in harmonizing access to hydrologic data through web services from university as well as federal data sources. There are presently over 5.2 billion data values from 1.96 million monitoring sites comprising 18,000+ variables in the federated system, and use has grown to an average of 7,000 web service data requests per day (doubling since 2010). To continue this success and support future growth in the sharing and collaborative use of hydrologic data and to satisfy the needs of the hydrologic community, a strategy for sustainable expansion of HIS is needed in three key areas: *interactive data access*, *model and tool sharing*, and *education*.

We propose to develop an **online, collaborative environment centered on the sharing of hydrologic data and models** that combines powerful, existing NSF funded technologies in a way that is sustainable for the water community. We refer to this new system as *HydroShare*. HydroShare will combine elements from three NSF-supported CI projects (CUAHSI HIS, HUBzero and iRODS). CUAHSI HIS will provide source material including data sources and modeling tools. HUBzero (McLennan and Kennell, 2010) will provide the core of a new collaboration environment for hydrology researchers. The integrated rule oriented data system (iRODS, Rajasekar et al., 2006) will provide an efficient and transparent data transport layer for collecting and translating large datasets for analysis and extending HydroShare to other communities. HydroShare will implement a web-based collaboration portal based upon HUBzero that allows users to share data and models, while extensions to the existing CUAHSI HIS HydroDesktop client will enable transparent sharing of results from desktop tools directly to HydroShare, with no user intervention required. This will give more hydrologists and water resources scientists' instant access to data and models by transparently handling data formatting and data transmission details. Our expectation is that this will enable better and faster capabilities for prediction, simulation and hypothesis testing necessary to advance hydrologic understanding and facilitate better water management.

HydroShare is based on the premise that a hydrologist should not need to understand technical details of a particular software infrastructure to be able to access and collaborate around data and models using HIS, or any CI. Consider how easily one can share and tag photos and other digital resources using a site such as facebook.com without having expertise in image data formats or metadata constructs. Such simple and functional collaborative systems do not currently exist for a scientist needing to share hydrologic data or tools with a colleague. At present, scientists can use CUAHSI HIS to establish a data server, load data into a database and make the data accessible using web services, but not every researcher wants to or has the ability to set up a HydroServer. With models, collaborative sharing is even more difficult with the potentially overwhelming requirements for documenting and communicating the functionality of models as well as the steps taken to parameterize, calibrate, and evaluate a modeling application. The result is that both model codes and model applications are frequently reinvented rather than built upon – slowing the progress of scientific advances. HydroShare addresses this problem by improving capability for model sharing, documentation, and reuse, including ease of sharing of both models and recipes for particular data products, as well as automatic, comprehensive record-keeping of how specific analyses were accomplished.

The current CUAHSI HIS is comprised of three elements: 1) data servers that store and publish hydrologic data (HydroServer); 2) desktop client software for data discovery, access and analysis (HydroDesktop); and 3) a centralized, ontology-based metadata catalog that integrates metadata from data servers to support discovery of data by the client (HydroCatalog) (Tarboton et al., 2009; 2010; CUAHSI, 2011). These components interoperate through web services that transmit information in an XML schema called WaterML (Zaslavsky et al., 2007), which is one of the foundations of WaterML 2.0, a new specification now being advanced through the Open Geospatial Consortium (OGC) standardization process as an international standard for communication of water data. The present HIS software development project ends this year, and CUAHSI is planning to operationalize the current software in a new data center to maintain data publication, discovery and transmission capabilities. This proposal will build on and extend these services. HydroShare will add a new class of functionality to HIS in the form of collaborative data and model sharing. This adapts the concepts of Web 2.0-style social networking to the sharing of hydrologic data and models. HydroShare will be a web application and as such will provide the benefits cross platform multi-user access with immediate version updates. In short, **the goal of HydroShare is to provide a software infrastructure that extends the data sharing capabilities of HIS and enables straightforward model and tool sharing.** Its collaborative environment will engage researchers and educators in knowledge creation that should enable development of new integrated models, regional and national scale water models, and better tools to advance understanding and lead to better water science and management.

A single platform for straightforwardly using, storing, and sharing both data and models has not been available to the hydrology community. Solutions to many of the important water-related problems facing society today require predictive capabilities for interacting processes over scales from less than a meter to the entire continent. The lack of ability to share models and their data is one of the reasons why answers to questions related to effects of climate change, or large-scale changes in agriculture for bio-fuels have been difficult to answer. This shortcoming has motivated discussions about developing community models, similar to the community models in other fields such as the atmospheric sciences (CUAHSI, 2009a; Famiglietti et al., 2010). HydroShare will provide critical CI required to support such a community hydrologic model.

2. New Interactive Data and Model Sharing for CUAHSI HIS

The present functionality and data cycle of HIS serves as a basis for illustrating functionality envisaged for HydroShare (Figure 1). Circled numbers in this figure refer to steps in the text that follows. Data are observed (1) and then loaded into HydroServer (Horsburgh et al., 2009; 2010) (2). HydroServer stores data in a relational database (Horsburgh et al., 2008) and publishes it using web services (Zaslavsky et al., 2007; Goodall et al., 2008) that make the data and metadata openly available using a services-oriented architecture. Metadata is harvested at HydroCatalog (3) into a global metadata repository that supports geographic and context based data discovery (Whitenack et al., 2008; 2010). A HydroDesktop (Ames et al., 2009) user discovers the data using search services supported at HydroCatalog (4) and then downloads the data from the HydroServer on which it is hosted directly into HydroDesktop (5). Key HIS contributions are the service-oriented architecture based on standards (Zaslavsky and Maidment, 2011), WaterML (Zaslavsky et al., 2007) for data transmission, Observations Data Model (ODM) (Horsburgh et al., 2008) for persistent data storage, and software at HydroServer, HydroDesktop and HydroCatalog.

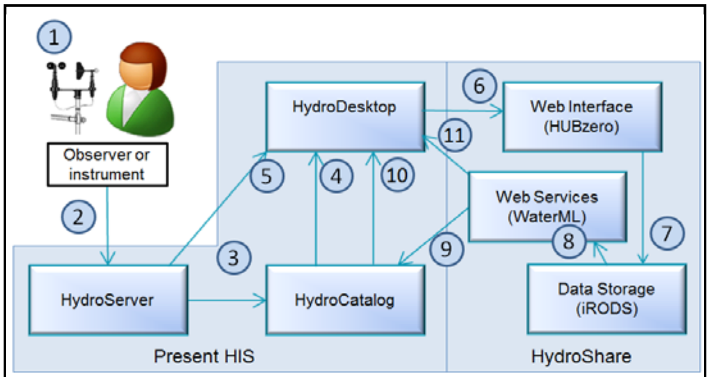


Figure 1 Data analysis and publication use case. Numbers refer to steps in the text

Presume that at this point in the HIS research cycle, a researcher selects data from HydroServers, filters the data in some way, and performs an analysis with a particular model resulting in some derived data or results – all of which need to be shared as part of a collaborative process. HIS was envisioned as a vehicle for capturing and publishing raw data, but does not have an easy way to facilitate sharing of models, or recording the metadata for model or analysis results placed in a HydroServer. HydroShare is envisioned as the solution. The user now posts the derived data and models to HydroShare using HydroDesktop (6). HydroDesktop ensures the data and models are in a format compliant with a community adopted standard, much of which HydroDesktop will have retained from the original data source and augmented based upon the analysis performed. Upon ingestion (7), iRODS micro-services parse the data, exposing the data and metadata through web services (8). Access control is set by the user, drawing upon functionality in HUBzero and iRODS. Metadata pertaining to the data is harvested by HydroCatalog (9). A collaborator discovers the derived data (10), perhaps having been alerted to its presence through the collaboration forum functionality implemented using HUBzero. A collaborator downloads the data using HydroDesktop (11) supported by new web services from HydroShare. At this point steps (6) through (11) may iterate as a group of collaborators perform their analyses, create key graphics, discuss major conclusions and outline a paper. Once the work is complete, the paper may be submitted for publication describing the final resulting datasets, which are assigned a unique identifier (e.g., digital object identifier) and made immutable and open to the public. Datasets would then become part of a curated body of knowledge following the emerging paradigm for publication of data as well as scientific writing (Hey et al., 2009). We will work with publishers such as the American Geophysical Union (see commitment letter) to explore ways for data and models published using HydroShare to extend and become part of their scientific publication system within this new paradigm.

A key point related to the process illustrated in Figure 1 is the interactive nature of the collaboration and sharing proposed. The illustration of Figure 1 focused on data. However many of the same ideas apply to models and methods encoded as scripts or workflows in some analysis systems. Collaborators need to discuss and interact around these too, and HydroShare will enhance this interaction capability.

Many of the capabilities outlined here are already present in technologies such as CUAHSI HIS, iRODS, and HUBzero, which we intend to exploit – thereby taking advantage of major NSF software development investments to create a complete digital hydrologic system. Our proposed system architecture (Figure 2) depicts how we plan to integrate these components.

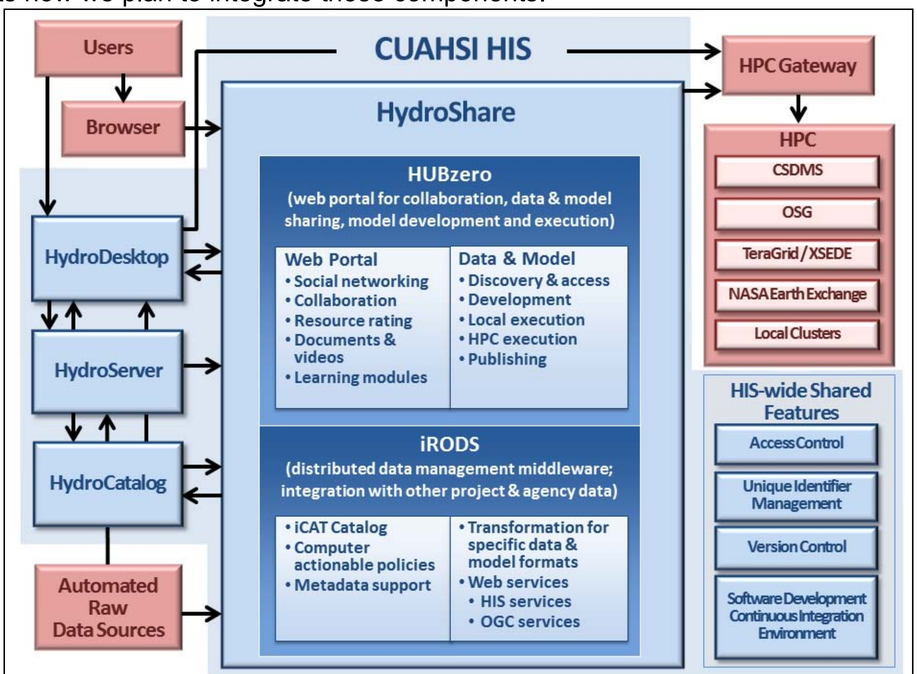


Figure 2 Proposed Conceptual Design

CUAHSI HIS currently has the ability to discover, access, analyze and publish observation time series, as well as publish spatial data. We will now add HydroShare to CUAHSI HIS to provide a web portal for collaboration, data & model sharing, and model development execution. HUBzero and iRODS will form the core of HydroShare. HUBzero will bring to HydroShare the platform for sharing organized collections of data and models, to set up interactive simulation sessions that access grid computing resources, and the ability to create and participate in user groups, which provide both a means of organizing people and content, as well as a mechanism for managing data access. We will build upon the NSF-funded open-source WaterHUB HUBzero implementation (Merwade et al., 2010) and include capability for sharing other online resources such as presentations, videos and documents that support educational uses. iRODS will serve as distributed data management middleware to support long-term preservation, and will provide the ability to federate HydroShare data with other community systems such as the Community Surface Dynamics Modeling System (CSDMS, 2010) and other agency systems such as NASA Earth Exchange (Nemani et al., 2011). In this regard, iRODS becomes an "on ramp" for HydroShare users into a national cyberinfrastructure.

The integration of all of the components depicted in Figure 2 to create HydroShare represents a major technological and enabling step for hydrology. HydroShare will be transformative in the capability it provides for integration of both spatial and temporal data, metadata, tools, and models. HydroShare will support a broad class of data types used in hydrologic analysis, extending the data types presently supported by CUAHSI HIS. New data types will include raster or vector geospatial data, gridded information (e.g., in GeoTIFF, or NetCDF format), and documents. The interfaces for interaction with HydroShare will be carefully specified, open, and, wherever possible, standards based to provide a mechanism for interaction with HydroShare from software and systems developed by others. In addition to directly hosted data, HydroShare will also include data in the form of references to external online data resources such as web services or web sites hosted by other parties so as to leave the data with its primary provider who is generally best suited to curate it. This follows the model presently used in HIS for data sources such as the USGS National Water Information System (NWIS) streamflow data where the USGS has established a WaterML-based GetValues service that provides access to data that it curates (USGS, 2007).

The resulting system will enable the collaborative assembly of data and models required to help realize CUAHSI's community vision of *developing an integrative understanding of interactions among water, earth, ecosystems, and society, and ultimately an ability to simulate water everywhere at all times.*

3. Community Driven Research and Development Agenda

As a community consortium, CUAHSI plays an important role in setting and articulating the community research agenda and advocating for the resources and facilities necessary to enable and carry out this research agenda. CUAHSI serves the community by: 1) developing, supporting, and operating research infrastructure; 2) improving and promoting access to data, information and models; 3) articulating and advocating priorities for community level water-related research and observations; 4) facilitating interactions among the diverse water research community; 5) promoting interdisciplinary education centered in water science; and 6) translating scientific advancements into effective tools for water management and policy (CUAHSI, 2010). CI plays a critical role in CUAHSI's community serving infrastructure, as it is through CI that holistic information is formed from the diversity of observations of water cycle processes (CUAHSI, 2002a; 2002b; 2007). CUAHSI surveys and needs assessments have guided prioritization of CI activities (Bandaragoda et al., 2005; 2006; CUAHSI, 2009b).

Shortcomings in the capabilities of current models have motivated recent workshops to evaluate community model development (CSDMS, 2009; CUAHSI, 2008; 2009a; U.S. EPA, 2008; Famiglietti et al., 2010). The consensus of these workshops is that in order to address modern science challenges, the community must come together to share resources and collaborate in ways that are beyond the scope of current paradigms. A stated goal of one of these efforts (CUAHSI, 2010) is coordinated, continental-scale modeling capability achieved through community modeling. The vision of such an activity is a harmonization of community, CI, and collaboration across natural and social science research boundaries to promote new discovery and address the larger water issues facing society within a community research and collaboration environment. HydroShare is critical CI required to support this vision.

Through the auspices of CUAHSI as a community consortium, the project team has the ability to engage and draw upon community input and user perspectives crucial to ensure adoption and sustainable success. The functionality proposed for HydroShare has been developed in response to a significant community need and will provide CI critical for integrating water data and models in a way that leads to better basic understanding of water processes and better solutions to critical water problems.

4. Integration of Existing Systems

In this section, we review the existing systems upon which HydroShare will be based. The architecture of the existing CUAHSI HIS is detailed in a series of reports and papers (Zaslavsky et al., 2007;

Maidment, 2008b; 2009; Horsburgh et al., 2008; 2010; Tarboton et al., 2009; 2010; Maidment et al., 2009) and is described in Section 2 and Figure 1. This proposal is to develop additional functionality comprising 1) simplified posting of data directly from HydroDesktop and other potential clients, 2) a separate central portal (HydroShare) where individual scientists (including those not associated with a HydroServer) can share data and models, and 3) an ability to federate community data with other community systems such as CSDMS and other agency systems such as NASA Earth Exchange (Nemani et al., 2011).

HydroShare will enable data and model sharing by adopting the open source HUBzero (McLennan and Kennell, 2010) collaboration and publishing framework. HUBZero allows scientists to collaborate on projects, upload and link data and papers form groups around data and models and support these groups with tools such as wikis, discussion forums, blogs, and wish lists for group members. HydroShare, when integrated with HUBzero, will provide a single hub for discovering, accessing, and publishing models and data, as well as running models and workflows on external shared resources such as the Open Science Grid (OSG) and TeraGrid. WaterHUB (Merwade et al., 2010) adds to HUBzero the capability of handling large spatial and temporal datasets with data-driven, computationally intensive hydrologic simulations.

Unlike the current HUBZero product, which supports only manually created associations and links between published objects, HydroShare will also create links transparently and without user intervention, by annotating data with the research processes that created them. This "data provenance" along with users' manual annotations will then serve as input to search services that can locate linked data and research results in a heretofore inconceivable way. For example, users will be able to search for academic papers that concern a specific watershed, and, by following the HUBzero link from paper to analysis results, and then from analysis results to raw data, the full workflow that resulted in the published paper will be available to HydroShare users. Further, HydroCatalog links from raw data to geospatial location, and from geospatial location a GIS server can link to any or multiple pertinent geographic layers (e.g. USGS Hydrologic Unit Code (HUC) watersheds), thus implementing **geo-centered publication search.**

HydroShare will gain data conversion and transport capabilities from iRODS (Rajasekar et al., 2006). iRODS offers flexibility for storing and working with data objects. It allows the data to be converted "on-the-fly" via computer-actionable usage policies, micro-services and workflows represented using iRODS rules (Moore et al., 2009). The iRODS architecture customizes data management functionality in an easy, declarative fashion in a service-oriented framework. An iRODS feature currently being developed in response to CUAHSI and other communities' needs will enable it to provide native iRODS integrated access to databases stored within iRODS. The new "database-objects" will virtualize database queries, allowing users to publish data comprised of many data sources such that the physical locations of data and also locations of data elements within specific distributed databases is abstracted by the iRODS virtualization.

HydroShare will leverage existing HPC model framework investments; in particular the NSF funded CSDMS (see commitment letter). CSDMS offers access to a collection of models and tools created by scientists with interests in modeling earth surface processes. Many of the models of interest to the water science community have already been indexed within CSDMS. The CSDMS Hydrology Research Focus Group chaired by co-PI Goodall has been tasked with prioritizing models for refactoring to operate within the CSDMS Community Modeling Toolkit (CMT). CMT was designed specifically to ease the transition of community models to a High Performance Computing (HPC) environment, and CSDMS maintains a 1,368-node cluster that is available free of charge to the CSDMS community. We have chosen to focus on CSDMS for a number of reasons, including our prior work on the CUAHSI HIS project exploring modeling frameworks (Goodall et al., 2010; Castronova and Goodall, 2010).

4.1 Use by Broader Communities

For HydroShare to have value that is transferable to other domains, we will pay particular attention to creating the incentive to use and add to the software and intellectual content. We will do this first by adherence to data exchange and format standards. Our work with WaterML (Zaslavsky et al., 2007) and the OGC (Co-PI Valentine co-leads OGC WaterML 2.0 development, as the co-chair of the WaterML 2 Standards Working Group) exemplifies this. We will also work to ensure that our software framework enables other communities to more easily integrate with and use the data generated and managed

HydroShare by adopting and working with data formats consistent with other closely aligned domains such as climate and atmospheric science (e.g., Unidata for which Tarboton serves on the policy committee), engineering, environmental and land/resource management.

We will also work to integrate HydroShare with NSF's ongoing DataNet (Sustainable Digital Data Preservation and Access Network Partners) program. Horsburgh is a member of the Core CI Team for the Data Observation Network for Earth (DataONE) project, which is developing a services-oriented-architecture for preserving and providing access to scientific data. DataONE consists of a distributed set of "Member Nodes" on which scientific data are published and Coordinating Nodes that provide data discovery and other services. We will implement the necessary web service interfaces for HydroShare to participate as a DataONE Member Node, enabling DataONE to index the contents of HydroShare and provide discovery and access to HydroShare resources through their existing infrastructure. Additionally, this linkage provides a mechanism for potentially replicating data from HydroShare to other DataONE Member Nodes, increasing the reliability of HydroShare's data archival capabilities. Additionally, RENCI is a partner and Goodall is co-PI on the DataNet Federation Consortium (DFC) (Moore and Rajasekar, 2010) DataNet proposal, which was recently recommended for funding. DFC plans to use iRODS to federate CUAHSI-catalogued data with that of several other communities (Cognitive Science; Oceanography; Biology; Social Science; Earth Systems; and Engineering) to provide an enabler to a "national data infrastructure." The iRODS federation mechanism enables integration of otherwise isolated data collections residing on separate iRODS datagrids to be made available to the user through a single virtualized presentation of the combined datagrids using one of iRODS many stand-alone clients or client interfaces. Similar to DataOne, we will implement necessary web service interfaces natively within iRODS micro-services and federate HydroShare's iRODS datagrid with DFC, providing common access to relevant hydrology collections spanning these communities.

To address broader multi-agency integration and leverage data and model resources assembled by NASA, we have forged a reciprocal shared-resource collaboration with the NASA Earth Exchange (NEX, Nemani et al., 2011) (see commitment letter). NASA image information, models and super-computers will be available to HydroShare users. An additional strategic advantage to collaboration with the NEX group is the existing implementation of data resources and workflows to support BIOME-BGC, a core model element of the RHESSys ecohydrology model that we propose to use as one of our modeling exemplars. This collaboration will extend capabilities for operating RHESSys over regional extents.

It is critical that we coordinate with agencies that collect hydrologic data and interface with their existing information systems. Specifically, NOAA, NWS Office of Hydrologic Development, U.S. Army Corps of Engineers, U.S. EPA, NASA, and USGS are keepers and providers of datasets needed by hydrologists for parts of the overall science of water. Maidment will use his extensive contacts with agencies and institutions to coordinate the development of HydroShare so that it evolves in harmony with the cyberinfrastructure plans of other organizations. He has been successful doing this with the current HIS project. For example, the USGS is already publishing time series data from the National Water Information System in CUAHSI's standard WaterML format. Maidment is currently working with the USGS Council for Data Integration to promote further integration between CUAHSI and USGS as the USGS expands its cadre of Web services for data publication.

5. Engineering Process

5.1 Software Engineering Process

We will use the Agile software development methodology (Schuh, 2004) and Scrum framework (Schwaber, 2004; Gorakavi, 2009) to coordinate the collaboration of our geographically distributed team. These practices, which have been in place and actively used by the iRODS@RENCI team that co-PI Idaszak oversees in the development of an enterprise version of iRODS (Brieger et al., 2011) are detailed in the Management and Coordination Plan that accompanies this proposal.

5.2 Testing Methodology

We will stress test HydroShare under different file count, file size, networking, and federated data and model use loads across a representative heterogeneous environment to ensure that the infrastructure on which HydroShare is implemented maintains robust functionality and performance. RENCI is implementing a heterogeneous distributed testing cluster tied to its Continuous Integration facility. This cluster will allow for testing of iRODS and other HydroShare components. Daily testing will automatically be launched using a queuing system that automatically retrieves the latest version of the source tree and runs it on the cluster nodes for testing various resource and federation configurations. Once a version of HydroShare is deployed, testing scripts will be automatically executed for various features, work and data loads, as well as client tools. These testing scripts will be generated from a testing matrix, where each dimension represents a possible change in the system: platform, network topology, server configuration, system features and select clients. Once testing has been completed, logs will be aggregated and provided as a report for tracking and analysis. To validate completeness of the testing code, coverage tools will be employed to determine untested portions of the code.

5.3 Community Process and Sustainability Plan

To gain community input on the direction and priorities of the project we will submit the project for regular review by the CUAHSI Informatics Standing Committee, which will serve as an Advisory Committee for the project. CUAHSI also has an HIS Users Committee that will provide more detailed feedback on functionality and use tests.

In addition to community feedback, HydroShare will be sustained through commitments from commercial partners. If the current success of CUAHSI-HIS is any indication, HydroShare will provide business opportunities and business intelligence for commercial software vendors. For example we have a letter of commitment from KISTERS Corporation to assign a product design specialist to work with us. KISTERS water information system products are used by government agencies, research units, and private organizations around the world. KISTERS is actively involved in developing or implementing open source standards within the OGC, and has included a WaterML service endpoint in their WISKI software. By participating in the development of HydroShare, KISTERS benefits from an expanding set of services. KISTERS can potentially sustain successful parts of the system based on their commercial value.

The evaluation metrics described in Section 7 will be used to evaluate every software release, both quantitatively and qualitatively to ensure that community expectations are met. By building on and contributing back to large open-source NSF supported projects, namely HUBzero and iRODS, we help sustain these projects and their availability for sustaining HydroShare after the proposal period has ended. The HUBzero capabilities of HydroShare will vastly increase the number of hydrology users that have access to shared hydrology community assets and allow them to collaborate, comment on, and rate user-provided products across the community. This will provide user-driven insight into community priorities, informing the HydroShare governance of development needs and direction. Co-PI Idaszak was recently co-PI and workshop report editor of a Scientific Software Innovation Institute (S2I2) workshop (Ahalt et al., 2010) specifically addressing software and community sustainability across all NSF-funded environmental observation communities and the role Open Source plays in community innovation and sustainability. Idaszak will leverage RENCI's industry partners, including Red Hat and others, in our SSI instilling sustainable best practices in open source, governance, community architecture, and software engineering and setting the stage for our SSI to tie into an eventual S2I2. Education, training, and outreach will be performed throughout the process as described in Section 8 and will be sustained after the project by CUAHSI. HydroShare hardware infrastructure is based largely on community-provided systems and will continue to be so. Finally, **released Software will be distributed under the New BSD Open Source license** (Open Source Initiative, 2008).

6. Project Plan

The research proposed here is a unique integration of cyberinfrastructure, community processes, software engineering, and multidisciplinary research. We will extend elements of HIS and add the HydroShare web based collaboration portal. The major computer science research objectives of HydroShare are to:

(1) Develop a method of applying software engineering principles and community input to create an online collaborative environment for users working with hydrologic data and models;
(2) Extend CUAHSI HIS by integrating iRODS and HUBzero to facilitate data and model sharing and model execution via HydroShare; and
(3) Provide a means for realizing an extensible national hydrologic information infrastructure integrated with OSG, TeraGrid, CSDMS, other agency systems, and local cluster resources.

6.1 Project Tasks

The tasks for accomplishing these objectives are given here, followed by a proposed timeline.

Task1: Establish Engineering Process and Infrastructure

The first task will be to establish the Agile teams, Scrum framework, Software Engineering, Data Engineering, Community Process and associated development and build-and-test environments described above. We will purchase hardware and stand up the www.hydroshare.org domain. Hardware hosting of HydroShare will be primarily at RENCI with modest development capability elsewhere.

Task 2: Develop a HUBzero Based Community Portal

This task will focus on building the HUBzero based community portal for HydroShare, involving both software development to adapt data and model publishing into HUBzero and content management to disseminate and engage the broader community. In order to publish datasets with metadata, additional tools and software will be developed to ingest particular kinds of metadata, e.g., supporting various standard-based templates. As the set of metadata for model publishing is developed, support for model publishing and sharing will also be developed.

HUBzero's support for predefined content categories (e.g., publications, online presentations, articles, downloads) and for content ranking and review comments will be used in HydroShare for community-based feedback, providing popularity based ratings and feedback on available data and models. We will work with the HUBzero Consortium to integrate the capability of ranking/reviewing/tagging of datasets and models that are published in a hub. Additionally, data and model resource discovery tools will be created for HydroShare. These tools will provide an easy to use interface with keyword search (or searching on metadata attributes). They will interface with iRODS for data access. These tools will build on HUBzero capabilities of securely transferring large data files from the server to users' local computers (laptop, desktop) and iRODS access. HydroShare will also import the data packages produced by HydroDesktop, as well as incorporate new HIS-wide shared functionality for access control, unique identifiers, and versioning. We plan to complete a proof-of-concept prototype and first beta release of the HydroShare HubZero based portal in May of year two of this project as described in Section 6.2.

Task 3: Add HydroShare sharing and data provenance features to HydroDesktop

Updates to HydroDesktop will integrate with HUBzero and store data packages and details of research and data provenance in real time, as researchers work with data and models. This will allow access and instantaneous sharing of data and results with other researchers from either the desktop or web browser. Provenance will include – for each model run – the model used and details of the input datasets. This will allow sharing of not just data, but also, the process by which data was generated. We will refactor the necessary components of HydroDesktop to achieve this, separating graphical user interface and computational functionality to enable logging of computational steps and to better support cross platform use. HydroDesktop will also be updated to incorporate the new HIS-wide functionality for access control, unique identifiers, and versioning.

Task 4: Design the User Experience

The user interface for HydroShare (both web-based and client-based via HydroDesktop) will allow information sharing among hydrologists at a level that was previously impossible. The interface will be designed to capture the relationships between digital objects automatically and without user intervention, as research and analysis progress. The system will store the **data provenance** of each data object, which is the recipe whereby that object was created, including references to the models and source data utilized to create it. Model runs will generate provenance data, which will follow a resulting object for its lifetime. The capture of provenance will be matched with collaboration tools that take advantage of this metadata and will give hydrology researchers the abilities to:

a) query for other researchers looking at the same data or utilizing the same models.

b) completely reproduce the computation of a co-worker that provided a result dataset without interacting with the co-worker who created it.

c) (long term, as the system is adopted) query the system for which publications study which watersheds (geo-located literature search).

Task 5: Advance Data Access by developing and using Standards-Compliant Interfaces

New data formats will be supported and new web service interfaces will be developed. We will extend web service capabilities beyond delivering point observation series as WaterML, to include standardized formats such as NetCDF, shapefiles, grids and other geospatial data formats as well as hydrology specific data formats, such as for river channel cross sections. Where required, iRODS micro-services will be developed to translate managed data into standard formats and services for use in models. These capabilities will enable HydroShare to provide access to data that provide a more complete picture of the hydrologic environment, beyond what time series data alone can do.

Implementing OGC standards will promote interoperability of HydroShare data with those from other communities and facilitate tighter links to existing GIS applications. We will implement OGC/WMO Hydrology Domain Working Group (HDWG) best practices and standards and applicable portions of the European INSPIRE standards for cataloging and distributing spatial information (European Commission, 2010). The OGC standards will include services for mapping and delivery of spatial data (WMS, WFS, WCS), cataloging (CSW), delivering observation data (SOS), and processing services (WPS). The experience of this project will, in turn, inform OGC/WMO HDWG best practices and WaterML development (Valentine is co-lead of OGC WaterML 2.0 development).

Under this task we will also develop data services to interoperate with DataOne and DFC DataNet projects as outlined above.

As an exemplar for advancing data access, we will establish a national repository within HydroShare for river channel cross section data: a new data type not presently supported by CUAHSI HIS, or the HDWG. Since 2003, the United States has spent more than $2 billion on digital flood map modernization. A great deal of river channel cross-section, morphology and hydraulic modeling data has been developed to support this mapping and some of that could be repurposed to advance water science. This repository will include a mechanism for voluntary submission of information and it will provide access to this data in a standard way such that it is easy to run hydraulic models that use this data on either local or HPC environments.

Task 6: Enhance Catalog Functionality and Add Standards-based Interfaces

HydroCatalog is the central metadata catalog for CUAHSI HIS and provides search services for HydroDesktop and other HIS clients. *Four HydroCatalog extensions will be implemented:* 1) handling data types other than point observation time series from CUAHSI water data services; 2) improving metadata harvesting by using a synchronization framework for catalog updates; 3) enabling iRODS iCAT capabilities for grid federation of data sources; and 4) adding functionality in support of social interaction, including data tagging and the ability to query for tags, access control, unique identifiers and versioning.

An iRODS iCAT-based HydroCatalog will allow users to register and manage web services and content from many data types (e.g., grid, NetCDF) as they are stored in iRODS. Enhanced discovery for HydroCatalog resources will be developed based on OGC Catalog Services for the Web (CSW). As part of this work we will prototype a CSW profile for managing time series records. This work will be shared with the HDWG and other OGC domain working groups for best practices and standardization.

Two new HydroCatalog web services will be added to enable social interaction and data sharing including a) the ability to “tag” datasets as relevant to a specific user or project; and b) the ability to query HydroCatalog for those datasets tagged as relevant to a specific user or project. These services will allow group vetting of datasets for validity and relevance through HydroShare client applications that use these services.

Task 7: Advance Model Access, Sharing and Communication

HydroShare will advance model access and sharing by developing metadata standards for models and a mechanism for sharing model codes and model applications. *We will develop a set of Model Metadata standards* sufficient to describe models and enable search and discovery of models. The model metadata will support three new types of information: (1) model data packages (which include model input and output data), (2) model codes, and (3) modeling tools. HydroShare will support the sharing of model data packages, which are specific instances of a model for a place, time, and scientific question. A broad range of models will be supported, including physical, empirical and statistical models. We will develop pre and post processing tools needed to transform raw data collected by in-situ and remote sensors to the specific input files required by a model and as assimilated data products for real-time modeling. We will enable users to share these modeling tools with colleagues through HydroShare. The metadata description for each of these cases will enable users to find models of interest from a collection of potentially thousands of models, model codes, and modeling tools that we anticipate that community members will upload into HydroShare. The metadata will include, but is not limited to, model attributes such as the processes represented, geographic extent and temporal domain of the model, the outputs and inputs of the model, and compatibility with a particular modeling framework (e.g., HydroDesktop, OpenMI, CSDMS, OMS) or a particular scripting language (e.g., R, Matlab, or Python).

To provide initial model content for HydroShare we will select and include key models from existing repositories of codes relevant to the water sciences (USGS, 2010; IGWMC, 2010; IEMSS, 2010; EDRC, 2010; CSDMS, 2010). While we will examine a wide set of model repositories for this work, our highest priority will be the CSDMS code repository. CSDMS, with NSF support, has indexed a large set of the hydrologic science models and tools used in the community and we will work with CSDMS (see commitment letter) to build interoperable CI for model sharing between HydroShare and CSDMS. We will build tools and the interface that gives users the ability to collaborate around these models through commenting, tagging, and rating modeling resources using enhanced HUBzero functionality as described in Task 2.

Task 8: Interface with Existing Environmental Observatories

We will use the Baltimore Ecosystem Study as a test bed for the integration of data and model sharing proposed here. This will demonstrate using HydroShare for conducting fundamental water science research, leveraging resources in NSF funded LTER, WSC, and CZO sites.

Our modeling will rely on the Community Modeling Toolkit (CMT) framework for model coupling. CSDMS is designed to ease access to High Performance Computing resources for earth scientists. This work will involve refactoring specific ecosystem and hydrologic models for inclusion in the CSDMS Community Modeling Toolkit (CMT). Our plans include refactoring RHESSys (Band et al., 1993; 2000; Tague and Band, 2004) and the Soil Water & Assessment Tool (SWAT, Arnold and Fohrer, 2005) models in order to combine the hydrologic and ecosystem strengths of RHESSys with the agricultural and water management strengths of SWAT. Band is an original developer of RHESSys and Goodall has current NSF support to refactor SWAT for component-based modeling and for operation within a cloud computing environment. These projects will help to inform the work required to move these models to CSDMS and a HPC paradigm. Both RHESSys and SWAT are listed in the CSDMS model repository as candidate models for refactoring to the CSDMS Initialize, Run, Finalize (IRF) protocol. By performing the refactoring work through this project, we will make RHESSys and SWAT components available to the larger CSDMS community for linking with other CSDMS models. In RHESSys, individual processes are written as separate components, attached to specific land surface classes in an object hierarchy. Model setup and parameterization includes use of advanced digital terrain analysis, remotely sensed canopy information, and soils, land cover and land management data sets. SWAT includes agricultural specific conceptualizations that are not present in RHESSys. By refactoring SWAT and RHESSys for CMT, it will be possible to couple more advanced agricultural components of SWAT with the ecohydrologic components in RHESSys and to use this model to test scientific hypotheses that span watersheds across a range of land use and management.

The Baltimore Ecosystem Study (BES) is an NSF Long Term Ecological Research (LTER) site, which was also funded as a Waters Testbed site and a NSF-WSC project. The site has an active CUAHSI HIS Server, and Co-PI Band is a founding member of the LTER, providing seamless access to diverse data sets including hydrologic, ecosystem, biogeochemical, land cover and social science information. Existing RHESSys models have been built for watersheds ranging from forest control catchments to suburban and denser urban sites. This provides a range of different conditions and primary information sources that dominate water, carbon and nitrogen cycling under conditions of significant urban growth, land use and demographic change, with consequences for downstream water bodies (water supply reservoirs and the Chesapeake Bay). This theme has been identified as one of the “Grand Challenges of Environmental Science” (NRC, 2001). A recent NRC report (NRC, 2011) has recommended the development of a “modeling institute” to facilitate scientific advances needed to manage the Chesapeake Bay and contributing watershed, by promoting improved informatics and simulation technologies.

We will use HydroShare to integrate and improve access to the considerable data resources available in the BES. Working with BES we will prototype and refine the CI developed for HydroShare, helping ensure its applicability in this real world research context. Prototyping in the first two project years will be extended for the next three years as tools for new fundamental science become available as part of the

Baltimore and Coweeta LTER (Band is co-PI on both projects) and with other selected NSF funded observatory sites (LTER, CZO) as part of existing inter-site comparison. The outcome of this comprehensive demonstration example will illustrate how the HydroShare system can be used to support hydrologic science and education.

Task 9: Develop User Support Materials and Outreach Workshops

One fundamental aspect of HydroShare will be to ensure usability and functionality that will ensure its early adoption and success. As a community consortium, CUAHSI provides a mechanism to engage users early and often in the process of developing and exposing HydroShare. CUAHSI has an active HIS Users Group, and Informatics standing committee, both of which will be utilized as points of contact with potential users of HydroShare to provide critical users' perspectives. We will develop extensive user materials such as functionality lists, online demonstrations, annotated videos and workbooks that demonstrate example use cases and important features of the web portal. Training materials that can be used by both instructors and researchers to learn how to use the HydroShare system together with HydroServer and HydroDesktop elements will be developed as part of CUAHSI's increasing library of HIS training materials. These materials will be developed in coordination with CUAHSI to support workshops and training sessions offered by CUAHSI personnel as described in the outreach section of this proposal.

Task 10: Facilitate and Ease Access to Use of High Performance Computing

We will augment CSDMS's HPC resources by developing the capability for alternative HPC job processing beyond the single cluster currently available. At present there is no specified provision for executing CMT workflows on other machines, or ability to take advantage of NSF hosted grid services such as Open Science Grid (OSG) or TeraGrid (e.g., Wilkins-Diehr et al., 2008; Foster, 2005; Frey et al., 2002). We will leverage HUBzero's capabilities to distribute HPC and large ensembles of serial and small-way parallel jobs to the TeraGrid, OSG, CSDMS, and to local campus resources by connecting the Hub into RENCI's EngageVO infrastructure (McGee et al., 2009) and utilizing HUBzero's Rappture toolkit. New tools can be constructed using the Rappture toolkit, which takes a description of the input/output data objects for a tool and generates a graphical user interface with integrated visualization capabilities. The existing Rappture toolkit works well in many scientific domains, but is currently lacking support for hydrologic models and data types. This task will investigate and develop software components to support the creation of hydrologic models to run in HUBzero and remotely on large computational resources such as TeraGrid and OSG. The WaterHUB project (Merwade et al., 2010) has *extended* HUBzero to support an online SWAT simulation system. This system allows users to upload, configure and run SWAT models on distributed computing resources in a web browser and to share, discover, and download existing SWAT models. We will extend this capability to RHESSys then build the componentized, integrated system to operate in this form. When a model is to be run, data sources will be prepared for the model, migrated to the resource, the model will be instantiated, and a model run will be triggered. iRods micro-services may be utilized to handle the data transformation and migration. When complete, results will be stored in iRods, and workflows to prepare the results for presentation via services will be executed, and the results will be registered. Provenance of the results run on HPC systems will be tracked, allowing for users to extract processing history from a dataset managed in HydroShare. We will add the capability to run large-scale models using HPC through HydroShare or through the HydroDesktop client software. We will also investigate how these implementations can support the models outlined for this project, as well as new software components needed to interface with scientific modeling frameworks such as those in CSDMS. This will result in the model workflows stored within HydroShare having access to significantly more compute capability via large scale high throughput computing models (McGee et al., 2009).

6.2 Proposed Project Timeline

The timeline is shown in Figure 3. A key principle of the Agile method that we will be using is recognizing that the complete product functional requirements cannot be fully anticipated. We have created a timeline that illustrates generally when the design, active development, and refining activities associated with our proposed tasks are anticipated. Development cycles will be short, allowing planning and implementation to be revisited often, building flexibility into the process and providing the ability to adapt and be responsive to changes in the timeline and associated technical requirements, community requests, staff availability, and disruptive technologies. In the Agile process the architecture and design of the system evolve during sprint development. In our case each sprint will consists of a three week iteration. Every six sprint cycles (i.e. approximately three times a year) in our Agile development process a version of HydroShare will be ready for distribution. Beginning in Year 2, these three releases per year will comprise two minor (i.e. beta) releases (May and September) and one major (i.e. "production") release (January) and will follow this rhythm Years 2 through 5. The NSF-required initial prototype and proof-of-concept will be associated with Task 2 representing the first beta release of HydroShare in May of Year 2. The success metrics defined in Section 7 will be measured on each beta-release and provided to the HydroShare Executive Committee. Bug testing, verification and validation, performance tests, and usability surveys will be performed via the HydroShare beta-site and, when completed, the beta-site prototype will be deployed to the production site for release. This procedure will be followed for each release. The "Refining" activity preceding tasks 5 and 6 reflects that these are based on ongoing HIS standards activities in progress prior to the commencement of this proposal.

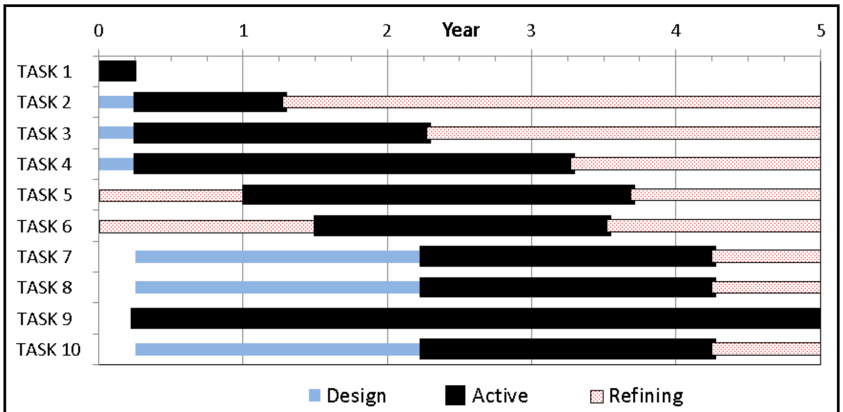


Figure 3: Proposed Timeline.

7. List of Metrics with End User Involvement

The success of HydroShare will be measured using metrics designed to evaluate its adoption and successful use by the hydrology community to do better science. The measure of success for any aspect of HydroShare will be determined by both the internal process used to create it, as well as by the community it serves. Internally we will follow well-established procedures in Agile / Scrum development as well as build and test methodologies. Each sprint and release milestone set forth will have iteratively established metrics to meet. Metrics will reflect measures for consistency in build and operation, timeliness in response to user interaction and sustainability in code reuse and function. Externally the project will initially adopt the following set of metrics noting that the selection and validation of metrics is an iterative process subject to community input and prioritization. Metrics will be tracked over time.

- The number of registered HydroShare users and number of host institutions of these users.
- Use statistics such as maximum and average number of users logged on, average log-on duration and total use time, total CPU hours of model run time, total run time using HPC.
- Geographic diversity of system users (based on IP address for casual non registered users).
- Number of courses and students using educational material.
- Diversity of user profiles accessing educational resources.
- Number of ratings and comments about resources.
- The quantity of hydrological data including data values, sites, and variables, and web service data requests per day.
- The number of non-CUAHSI agencies that utilize HydroShare (e.g., NCDC). HydroShare enables other agencies to more readily take advantage of the models and data available.
- The number of contributors to the code, both to the core and non-core portions of the repository.
- The number of downloads of each subsequent release of the code.
- The number of users trained during the various outreach activities.
- New science generated. Measuring this will include the number of papers submitted to and published in peer reviewed forums about this project and its implementation distinguishing and measuring both domain research papers from infrastructure papers. We will also measure posters, invited talks, panel sessions, etc. We will also track citations generated by these papers.
- Number of citations of various HydroShare resources.
- The types and amounts of data stored within the system, and their associated downloads.
- The number of models published within the system, and the number of model downloads.

8. Education and Outreach plan

The proposed elements of HydroShare will have a broad and diverse impact on learning that ranges from enabling novel functionality in the classroom to an innovative portfolio for independent learning and professional development. We anticipate HydroShare users from young students to full professors, and from practitioners to policy makers. Our education and outreach plan must, therefore, be broad and will focus on deepening understanding of hydrologic science by encouraging greater use of active, data-intensive approaches to learning and knowledge generation. *We hypothesize that the online HydroShare collaborative environment will facilitate active learning from introductory to advanced levels and that the time required to acquire critical modeling and data-intensive analysis skill sets will be significantly reduced because HydroShare will shift the effort from the mechanics of configuring models or manipulating data to using them to explore and develop deeper understanding of scientific principles and hydrologic processes*. We will develop a formal assessment methodology for each activity described below to gauge its effectiveness in enhancing learning and altering attitudes and perceptions about hydrology. To evaluate the effectiveness of the educational material, several dependent measures such as the concept verification task (CVT) and concepts application and transfer task (CATT) will be collected in the form of pretest and posttest surveys.

Design of the education and outreach activities will be led by Arrigo. CUAHSI staff (Arrigo and Choi) will implement these activities across willing CUAHSI member institutions. Several CUAHSI members are minority serving and have K-12 outreach programs that we will connect with to expose HydroShare to students as a way to stimulate interest in STEM. We will also collaborate with other projects having educational missions related to hydrology (e.g. MOCHA, 2010; CSDMS, 2010) by sharing and cross-listing resources to minimize redundancy and maximize community access to educational resources.

8.1 Novel classroom functionality

The access to models and data that HydroShare will provide will enhance instruction and learning across a wide range of natural science and engineering courses. We anticipate that with wide adoption of HydroShare, the user community will be a rich source of new educational uses, and foster educational and teaching collaborations in addition to research collaborations. Here we briefly highlight through examples how key aspects of HydroShare can be used to enhance learning.

HydroShare's simple online user interface and access to data can enable data-driven exploration and learning. Coupled with HydroDesktop, HydroShare can increase instructors' access to high quality, locally relevant environmental data that can be used in classroom applications, and support data-driven learning. We will develop short tutorials for educators on how to locate specific types of environmental data, and simple learning exercises for classroom use. For example, a short tutorial in the use of HydroDesktop could demonstrate how to obtain streamflow data over an area of interest, how to delineate a watershed, and how to find accompanying rainfall and/or temperature data. Either as a demonstration or as a class exercises, simple graphing and analysis of the data directly in HydroDesktop could be the basis for a discussion of concepts such as runoff, event flow, floods or droughts. Using collaboration tools in HydroShare, a class could collect subsequent water quality or quantity data over the same area (an exercise common in many environmental science courses) and upload it to HydroShare. This might enable collaboration with other classes in other locations, or provide a way for instructors to carry experiments or data exercises through subsequent semesters, courses or years.

HydroShare will provide access to models and data for use in classes where modeling and data-intensive analysis are part of the syllabus, and by so doing will support active problem-based learning that is the foundation of much modern pedagogy (e.g., Felder and Brent, 1999; Duch et al., 2001; King and Kitchener, 1994; Novak and Patterson, 1998; Riel, 1998). HydroShare will enable incorporation of more contextualized content in engineering programs that education studies suggest improve student learning, particularly that of women (Hayes, 2000; Daudt and Salgado, 2005; Du and Kolmos, 2009). As an example, students in a hydrology class would initially work with calibrated models representing real locations from a suite of models available in HydroShare. Students would explore problems using these models by posing hypotheses and testing them with numerical experiments involving altering a parameter, boundary condition or geometry and evaluating the resulting change in the model output. This type of exercise would train hydrologic intuition and hone problem solving skills. The next step would be for students to develop their own models and then conduct numerical experiments to learn more about the models and the processes they represent. Students who are having problems can readily get help by sharing their model with their instructor, colleagues, or potentially other on-line participants within HydroShare. Sharing of existing models and associated input/output and metadata including the training modules for the model will also overcome the steep learning curve barrier to some extent because the users already will have a base model to begin with, and they can spend more time on testing hypotheses and interpreting model results.

Overall, CUAHSI staff will lead the development of example data demonstrations, learning exercises, example data analyses and models, along with teaching webpages and tutorials designed to communicate the examples to educators and students. These examples are meant to "seed" the development of further educational materials and collaborations. We will enable the user community to submit examples of their classroom applications of HydroShare, broadening available resources and encouraging active participation from users.

Assessment of the educational activities will include tracking the usage of project-produced educational resources and the submission rate of new materials. Direct assessment of the project-produced examples will be done by introducing HydroShare in hydrology courses at the collaborating institutions, and conducting pre-and post course assessments of student learning using vetted online assessment tools (e.g. SALG, 2011). Results from surveys at collaborating institutions will then be used to refine and improve HydroShare as a learning tool and its overall usability. This refined HydroShare educational material will then be made public to CUAHSI institutions for adoption in appropriate courses. The open online availability of learning material will also make it available to independent learners.

8.2 Training in the Use of HydroShare

Many users may benefit from training on HydroShare. We will develop an online course and hold workshops to meet this need. The online training will build on the course materials developed by Whiteaker, Choi, and Horsburgh for the current HIS by incorporating instructional and community-oriented features from HydroShare, and by developing example uses cases and materials for specific audiences (e.g. researcher, modeling students, educator). Training workshops will be developed and conducted by the CUAHSI User Support Specialist (Choi), who routinely gives training workshops on CUAHSI HIS. The User Support Specialist will also provide ad hoc assistance to HydroShare users.

9. Summary of Broader Impacts

Broader Impacts of the proposed project range from advancing understanding and discovery to improving the training and education within water-related science and engineering. The proposed CI will catalyze new science by providing important new tools for answering questions related to climate change, land use change, bio fuels, carbon sequestration, and other problems critical to society. HydroShare will provide the CI foundation for linking together data and models in a National Water Model, which will usher in a new generation of simulation capability that will support science, management, policy, and decision-making related to water. The project will reach beyond academia and include informal collaborators and other investigators in federal agencies involved with data publication and modeling. CUAHSI has demonstrated the ability to engage agency personnel, and we will be building upon established relationships as we move forward. The online publishing of data and models will be a major factor driving broad impacts, and our collaboration with AGU will provide important leverage in the rapidly evolving online publishing arena. Our interactions with NSF's DataNet program via DataONE and DFC will broaden the impact of HydroShare data resources by making them readily available to other domains via interoperable formats and interfaces. Hydroshare's use and implementation Open Geospatial Consortium standards will expand the discovery and access to outside the community, and will expose the project to an international community. Feedback on usage of these standards will further development of OGC Hydrology Domain Working Group best practices.

HydroShare will build upon the broad impact of the existing HIS. The proposed functionality and our designs address needs that have been elicited in community surveys and workshops. The community governance involving 125+ CUAHSI member universities, as well as education and outreach activities supported by CUAHSI core office staff, will ensure the broad impact of this project across the hydrologic community. Our reliance upon established open source software, our adherence to existing data and metadata standards (and where no standards exist our participation in creation of standards), and the accessibility of our own software will facilitate broad distribution and impact.

The education and outreach activities have been designed to meet identified community needs and support active learning across all educational levels. Our education and outreach plan is ambitious and will provide materials for changing the way hydrologic science is taught. HydroShare will also be the focal point of a broad spectrum of live educational material that will be readily accessible on the Internet to anyone from anywhere. For example, a professor at MIT could learn how to use a new method for upscaling, a student at Utah State University could learn how to predict changes in water supply due to changing land use, or a water manager in California may learn how to improve forecasts of snowpack in a changing climate. The educational impact will be more immediate for a handful of students involved directly in the proposed project. We will support graduate and undergraduate students, who will participate in the development and refinement of HydroShare, contributing to the training of the next generation of scientists and software developers. Another, much larger, cadre of students will benefit from the enhanced learning opportunities available as HydroShare becomes adopted for classroom applications in the U.S. and abroad.

Results will be published in the peer reviewed literature as well as broadly disseminated at conferences, workshops and meetings. Finally, HydroShare will have broad impact by serving as a blueprint for combination of software tools and collaboration applicable to other disciplines.

10. Results from Prior NSF Support

GeoInformatics: CUAHSI Hydrologic Information Systems, PIs: D.R. Maidment, ***D.G. Tarboton***, I. Zaslavsky, ***J. Goodall***, and ***D.P. Ames***, EAR 0413265 and EAR 0622374, $1,156,059, 4/1/04 – 3/1/08 and $4,500,000, 1/15/07 – 1/15/12. The CUAHSI HIS project and Cyberinfrastructure serves as a foundation for the work proposed here. Publications: Ames et al. (2009), Beran and Piasecki (2008; 2009) Beran et al. (2009), Castronova and Goodall (2010), Goodall et al. (2008), Horsburgh et al. (2008; 2009; 2010), Maidment et al. (2006; 2009), Maidment (2008a; 2008b; 2009), Piasecki and Beran (2009),Tarboton et al. (2009), Zaslavsky and Maidment (2011), Zaslavsky et al. (2007).

Collaborative Research: CI-TEAM Demonstration Project: WaterHUB for Cyber Enabled Training, Education and Research in Water Resources, PIs: **V. Merwade, C. Song,** S. Brophy, B. Ruddell, and A. Yerrammilli, OCI 1041379, $ 222,107 09/01/10 – 08/31/12. The objective of this project is to develop a HUBZero based cyberinfrastructure called WaterHUB for sharing hydrologic data and modeling tools for training, education and research in Water Resources (Merwade et al., 2010).

NSF Workshop on a Scientific Software Innovation Institute for the Environmental Observatory Communities, PIs: S. Ahalt, B. Minsker, **R. Idaszak**, OCI 1049273, $50,000 10/1/10 – 3/31/11. **Co-PI Idaszak** was workshop host and subsequent workshop report editor. The workshop revealed the critical necessity of an S2I2 Institute to deliver a cyberinfrastructure software ecosystem addressing the grand challenges of environmental observation and related communities (Ahalt et al., 2010).

Human settlements as ecosystems: Metropolitan Baltimore from 1790- 2100, **Larry Band** is one of 13 PIs, 2004-2011, DEB 0423476 NSF Long Term Biology, LTER Network, $7,200,000 (UNC portion-$390,000) Three graduate students supported, one masters thesis, two PhDs, eight published journal articles (Tenenbaum et al., 2006b; Tenenbaum et al., 2006a; Tague et al., 2006; Groffman et al., 2006; Smith et al., 2010; Kaushal et al., 2005; Kaushal et al., 2008; Pickett et al., 2008), three book chapters (Band and Tague, 2005; Band et al., 2005; Pouyat et al., 2007) focusing on urban-to-rural gradients of water, carbon and nutrient cycling and export. (renewed 2011-2017)

TeraGrid Resource Partner Dr. Carol Song is the PI for NSF award OCI-0503992, $5.9M, 8/2005–7/2013. For this grant, Purdue provides high performance and high throughput computation, data and storage, science gateway resources to the nation’s research and education communities. Selected publications (Katz et al., 2010; Braun et al., 2010; Carlyle et al., 2010; Zhang et al., 2010; Zhao et al., 2010; Zhao et al., 2009; Chuang et al., 2009; Younts et al., 2009; Smith et al., 2008; Sundaram et al., 2008; Basumallik et al., 2007; Zhao et al., 2006; Zhao et al., 2007).

References

Ahalt, S., B. Minsker and R. Idaszak, (2010), "NSF Workshop Report on a Scientific Software Innovation Institute for Environmental Observatories," NSF Workshop Report funded under NSF grant # 1049273, October 4-5, http://www.acsu.buffalo.edu/~abani/RENCI_NCSA_NSF_1049273_workshop_report.pdf.

Ames, D. P., J. Horsburgh, J. Goodall, T. Whiteaker, D. Tarboton and D. Maidment, (2009), "Introducing the Open Source CUAHSI Hydrologic Information System Desktop Application (HIS Desktop)," 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, ed. R. S. Anderssen, R. D. Braddock and L. T. H. Newham, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009, 4353-4359, http://www.mssanz.org.au/modsim09/J4/ames.pdf.

Arnold, J. G. and N. Fohrer, (2005), "SWAT2000: current capabilities and research opportunities in applied watershed modelling," Hydrological Processes, 19(3): 563-572, http://dx.doi.org/10.1002/hyp.5611.

Band, L. E., M. Cadenasso, S. Grimmond, M. Grove and S. T. Pickett, (2005), "Heterogeneity in Urban Ecosystems: Pattern and Process," Ch.13 in Ecosystem Function in Heterogeneous Landscapes, Edited by G. M. Lovett, C. G. Jones, M. G. Turner and K. C. Weathers, Springer-Verlag, NY.

Band, L. E., P. Patterson, R. Nemani and S. W. Running, (1993), "Forest Ecosystem Processes at the Watershed Scale: Incorporating Hillslope Hydrology," Agricultural and Forest Meteorology, 63: 93-126.

Band, L. E. and C. Tague, (2005), "Feedbacks and Coupling between Water, Carbon and Nutrient Cycling at the Hillslope Scale," Ch. 4.10 in Coupled Models for the Hydrological Cycle - Integrating Atmosphere, Biosphere, and Pedosphere, Edited by A. Bronstert, J. Carrera, P. Kabat and S. Lütkemeier, Springer-Verlag.

Band, L. E., C. L. Tague, S. E. Brun, D. E. Tenenbaum and R. A. Fernandes, (2000), "Modelling Watersheds as Spatial Object Hierarchies: Structure and Dynamics," Transactions in GIS, 4(3): 181-196, http://dx.doi.org/10.1111/1467-9671.00048.

Bandaragoda, C., D. G. Tarboton and D. R. Maidment, (2006), "Hydrology's effort towards the Cyberfrontier," EOS, 87(1): 2, 6, 10.1029/2006EO010005.

Bandaragoda, C. J., D. G. Tarboton and D. R. Maidment, (2005), "User Needs Assessment, Chapter 4," in Hydrologic Information System Status Report, Version 1, Edited by D. R. Maidment, p.48-87, http://www.cuahsi.org/docs/HISStatusSept15.pdf.

Basumallik, A., L. Zhao, C. X. Song, R. L. Sriver and M. Huber, (2007), "A Community Climate System Modeling Portal for the TeraGrid," TeraGrid 2007 Conference, Madison, Wisconsin, June 4-8.

Beran, B., J. Goodall, D. Valentine, I. Zaslavsky and M. Piasecki, (2009), "Standardizing Access to Hydrologic Data Repositories through Web Services," Proc. International Conference on Advanced Geographic Information Systems & Web Services (GEOWS 2009), Los Alamitos, CA, Feb. 2009, IEEE Computer Society, 64-67.

Beran, B. and M. Piasecki, (2008), "Availability and coverage of hydrologic data in the US geological survey National Water Information System (NWIS) and US Environmental Protection Agency Storage and Retrieval System (STORET)," Earth Science Informatics, 1(3): 119-129, http://dx.doi.org/10.1007/s12145-008-0015-2.

Beran, B. and M. Piasecki, (2009), "Engineering new paths to water data," Computers & Geosciences, 35(4): 753-760, http://dx.doi.org/10.1016/j.cageo.2008.02.017.

Braun, D., C. Thompson, A. Younts, P. Smith and X. C. Song, (2010), "HUBzero + Cloud: Power tool for the masses," 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2010), Indianapolis, IN, Nov. 30 – Dec. 3.

Brieger, L., J. Coposky, V. Dantuluri, K. Gamiel, R. Idaszak, O. Kapeljushnik, N. Nassar, J. Reilly, M. Stealey, L. Stillwell and X. Wang, (2011), "iRODS@RENCI," iRODS User Meeting 2011, February 17-18, http://iren-web.renci.org/irods-meeting/irods@renci-2011UserMeeting-contribution.pdf, accessed 6/24/2011.

Carlyle, A., S. Harrell and P. Smith, (2010), "Cost-effective HPC: The Community or the Cloud?," 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2010), Indianapolis, IN, Nov. 30 – Dec. 3.

Castronova, A. M. and J. L. Goodall, (2010), "A generic approach for developing process-level hydrologic modeling components," Environmental Modelling & Software, 25(7): 819-825, http://dx.doi.org/10.1016/j.envsoft.2010.01.003.
Chuang, H. C., D. Niyogi, L. Zhao, H. I. Chang, X. C. Song, M. Lei, R. S. Govindaraju, F. Chen and U. Charusambot, (2009), " A TeraGrid Science Gateway for Studying Hydrological Extremes using a Land Data Assimilation System," TeraGrid 2009 conference, Arlington, VA, June.
CSDMS, (2009), Hydrology Focus Working Group meeting January 20-21, 2009, http://csdms.colorado.edu/wiki/Hydrology_FRG_2009, accessed 6/5/2010.
CSDMS, (2010), Community Surface Dynamics Modeling System, http://csdms.colorado.edu/wiki/, accessed 6/13/2010.
CUAHSI, (2002a), "CUAHSI Hydrologic Information Systems," Consortium of Universities for the Advancement of Hydrologic Science, Inc. Technical Report Number 2 - Hydrologic Information Systems Committee, http://www.cuahsi.org/docs/dois/CUAHSI-TR2.pdf.
CUAHSI, (2002b), "A Vision for Hydrologic Science Research," Consortium of Universities for the Advancement of Hydrologic Sciences Inc., Technical Report Number 1, http://www.cuahsi.org/publications/cuahsi_tech_rpt_1.pdf.
CUAHSI, (2007), "Hydrology of a Dynamic Earth: A Decadal Research Plan for Hydrologic Science," Consortium of Universities for the Advancement of Hydrologic Sciences Inc. Science Plan, http://www.cuahsi.org/SciPlan-20070402.html.
CUAHSI, (2008), Scoping Workshop on a Community Hydrologic Modeling Platform (CHyMP), March 26-27, 2008, http://www.cuahsi.org/chymp-20080326.html, accessed 6/5/2010.
CUAHSI, (2009a), 2nd Workshop on a Community Hydrologic Modeling Platform, March 31 – April 1, 2009, http://www.cuahsi.org/chymp-agenda-20090331.html, accessed 6/5/2010.
CUAHSI, (2009b), HydroHUB: Proposed Modeling Infrastructure for the Water Science and Engineering Communities, http://www.cuahsi.org/hydrohub/, accessed 6/6/2010.
CUAHSI, (2010), "Water in a Dynamic Planet, A Five-year Strategic Plan for Water Science," Consortium of Universities for the Advancement of Hydrologic Sciences Inc., http://dx.doi.org/10.4211/stratplan.201012.
CUAHSI, (2011), CUAHSI HIS Sharing Hydrologic Data, http://his.cuahsi.org/, accessed 7/7/2010.
DataONE Project Team, (2011), DataONE structure and potential partnership as a Member Node, http://www.dataone.org/content/dataone-structure-and-potential-partnership-member-node, accessed 3/13/11.
Daudt, J. and P. P. Salgado, (2005), "Creating a woman friendly culture in institutes of higher engineering education," European Journal of Engineering Education, 30(4): 463 - 468, http://www.informaworld.com/10.1080/03043790500213060.
de la Beaujardiere, J., ed. (2006), OpenGIS Web Map Server Implementation Specification, OGC Implementation Specification OGC 06-042, Version 1.3.0, http://portal.opengeospatial.org/files/?artifact_id=14416.
Du, X. and A. Kolmos, (2009), "Increasing the diversity of engineering education: a gender analysis in a PBL context," European Journal of Engineering Education, 34(5): 425-437, http://www.informaworld.com/10.1080/03043790903137577.
Duch, B. J., S. E. Groh and D. E. Allen, ed. (2001), The Power of Problem-Based Learning, Stylus, Sterling, VA, 256 p.
EDRC, (2010), Engineer Research and Development Center, http://www.erdc.usace.army.mil/, accessed 6/13/2010.
European Commission, (2010), Inspire Geoportal, http://www.inspire-geoportal.eu/, accessed 7/15/2011.
Famiglietti, J., L. Murdoch, V. Lakshmi and R. Hooper, (2010), "Rationale and Strategy for a Community Modeling Platform in the Hydrologic Sciences, Report of the CHyMP Scoping Workshop, March 26-27, 2008 ", CUAHSI Technical Report #8, Washington, DC, http://dx.doi.org/10.4211/techrpts.200911.tr8.
Felder, R. M. and R. Brent, (1999), "Exceed Effective College Teaching Seminar," 1999 Civil Engineering Conference and Exposition, Charlottville, North Carolina, ASCE.
Foster, I., (2005), "Service-Oriented Science," Science, 308(5723): 814-817, http://dx.doi.org/10.1126/science.1110411.
Frey, J., T. Tannenbaum, M. Livny, I. Foster and S. Tuecke, (2002), "Condor-G: A Computation Management Agent for Multi-Institutional Grids," Cluster Computing, 5(3): 237-246, http://dx.doi.org/10.1023/A:1015617019423.

Goodall, J. L., M. Elag and A. M. Castronova, (2010), "Component-Based Modeling of Watershed Systems," ASCE Watershed Management Conference, Madison, WI.
Goodall, J. L., J. S. Horsburgh, T. L. Whiteaker, D. R. Maidment and I. Zaslavsky, (2008), "A first approach to web services for the National Water Information System," Environmental Modelling & Software, 23(4): 404-411, doi:10.1016/j.envsoft.2007.01.005.
Gorakavi, P. K., (2009), Build Your Project Using Scrum Methodology, http://www.asapm.org/asapmag/articles/A3_AboutScrum.pdf, accessed 6/6/2010.
Groffman, P. M., R. V. Pouyat, M. L. Cadenasso, W. C. Zipperer, K. Szlavecz, I. D. Yesilonis, L. E. Band and G. S. Brush, (2006), "Land use context and natural soil controls on plant community composition and soil nitrogen and carbon dynamics in urban and rural forests," Forest Ecology and Management, 236(2-3): 177-192, http://www.sciencedirect.com/science/article/pii/S0378112706008747.
Hayes, E., (2000), "Social contexts," in Women as learners – the significance of gender in adult learning, Edited by E. Hayes and D. F. Daniele, Jossey-Bass, San Francisco, p.23–52.
Hey, T., S. Tansley and K. Tolle, (2009), The Fourth Paradigm, Data-Intensive Scientific Discovery, Microsoft Research, Redmond, Washington, 283 p, http://research.microsoft.com/en-us/collaboration/fourthparadigm/.
Horsburgh, J. S., D. G. Tarboton, D. R. Maidment and I. Zaslavsky, (2008), "A Relational Model for Environmental and Water Resources Data," Water Resour. Res., 44: W05406, doi:10.1029/2007WR006392.
Horsburgh, J. S., D. G. Tarboton, M. Piasecki, D. R. Maidment, I. Zaslavsky, D. Valentine and T.Whitenack, (2009), "An integrated system for publishing environmental observations data," Environmental Modelling & Software, 24(8): 879-888, http://dx.doi.org/10.1016/j.envsoft.2009.01.002.
Horsburgh, J. S., D. G. Tarboton, K. A. T. Schreuders, D. R. Maidment, I. Zaslavsky and D. Valentine, (2010), "Hydroserver: A Platform for Publishing Space-Time Hydrologic Datasets," 2010 AWRA Spring Specialty Conference Geographic Information Systems (GIS) and Water Resources VI, Orlando Florida, American Water Resources Association, Middleburg, Virginia, TPS-10-1, http://www.awra.org/orlando2010/doc/abs/JefferyHorsburgh_7cb420e3_6602.pdf.
IEMSS, (2010), International Environmental Modeling & Software Society, http://www.iemss.org/society/, accessed 6/13/2010.
IGWMC, (2010), International Ground Water Modeling Center Alphabetic Listing of Programs, http://igwmc.mines.edu/software/igwmcsoft/, accessed 6/13/2010.
Katz, D. S., S. Callaghan, R. Harkness, S. Jha, K. Kurowski, S. Manos, S. Pamidighantam, M. Pierce, B. Plale, C. Song and J. Towns, (2010), "Science on the TeraGrid," Computational Methods in Science and Technology: 81-97, http://www.man.poznan.pl/cmst/2010/_SI_i3_2010/08_Katz_fin.pdf.
Kaushal, S. S., P. M. Groffman, L. E. Band, C. A. Shields, R. P. Morgan, M. A. Palmer, K. T. Belt, C. M. Swan, S. E. G. Findlay and G. T. Fisher, (2008), "Interaction between Urbanization and Climate Variability Amplifies Watershed Nitrate Export in Maryland," Environmental Science & Technology, 42(16): 5872-5878, http://dx.doi.org/10.1021/es800264f.
Kaushal, S. S., P. M. Groffman, G. E. Likens, K. T. Belt, W. P. Stack, V. R. Kelly, L. E. Band and G. T. Fisher, (2005), "Increased salinization of fresh water in the northeastern United States," Proceedings of the National Academy of Sciences of the United States of America, 102(38): 13517-13520, http://www.pnas.org/content/102/38/13517.abstract.
King, P. M. and K. S. Kitchener, (1994), Developing reflective judgement: Understanding and promoting intellectual growth and critical thinking in adolescents and adults, Jossey-Bass Publishers, San Francisco, CA.
Maidment, D. R., (2008a), "Bringing Water Data Together," Journal of Water Resources Planning and Management, 134(2): 95-96, http://link.aip.org/link/?QWR/134/95/1.
Maidment, D. R., ed. (2008b), CUAHSI Hydrologic Information System: Overview of Version 1.1, Consortium of Universities for the Advancement of Hydrologic Science, Inc, Washington, DC, 96 p, http://his.cuahsi.org/documents/HISOverview.pdf.
Maidment, D. R., ed. (2009), CUAHSI Hydrologic Information System: 2009 Status Report, Consortium of Universities for the Advancement of Hydrologic Science, Inc, 79 p, http://his.cuahsi.org/documents/HISOverview_2009.pdf.
Maidment, D. R., R. P. Hooper, D. G. Tarboton and I. Zaslavsky, (2009), "Accessing and Sharing Data Using CUAHSI Water Data Services," in Hydroinformatics in Hydrology, Hydrogeology and Water Resources, Edited by I. Cluckie, Y. Chen, V. Babovic, L. Konikow, A. Mynett, S. Demuth and D. A.

Savic, Proceedings of Symposium JS4 held in Hyderabad, India, September 2009, IAHS Publ. 331, Hyderabad, India, p.213-223, http://iahs.info/redbooks/331.htm.
Maidment, D. R., I. Zaslavsky and J. S. Horsburgh, (2006), "Hydrologic Data Access Using Web Services," Southwest Hydrology, 5(3): 16-17.
McGee, J., J. Reilly and M. Rynge, (2009), "Building Solutions for Researchers with a Science Gateway," Teragrid09, Arlington, Virginia, June 22-25, http://archive.teragrid.org/tg09/files/tg09_submission_75.pdf.
McLennan, M. and R. Kennell, (2010), "HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering," Computing in Science & Engineering, 12(2): 48-53, http://dx.doi.org/10.1109/MCSE.2010.41.
Merwade, V., B. Ruddell, C. Song, S. Brophy, R. Mohtar and A. Yerrrammilli, (2010), "Water-HUB - A community cyberinfrastructure for hydrology education and research," Abstract H43B-1238 presented at 2010 Fall Meeting, AGU, San Francisco, Calif., 13-17 December.
MOCHA, (2010), Modular Curriculum for Hydrologic Advancement, http://www.mocha.psu.edu, accessed 6/13/2010.
Moore, R. and A. Rajasekar, (2010), "iRODS: Data Sharing Technology Integrating Communities of Practice," Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, July 25.
Moore, R., A. Rajasekar, M. Wan and W. Schroeder, (2009), "Policy-based Distributed Data Management Systems," Policy-based Distributed Data Management Systems, Atlanta, GA, May 18-21.
National Research Council Committee on Grand Challenges in Environmental Sciences, (2001), Grand Challenges in Environmental Sciences, Oversight Commission for the Committee on Grand Challenges in Environmental Sciences, National Academy Press, Washington, DC.
National Research Council Committee on the Evaluation of Chesapeake Bay Program Implementation for Nutrient Reduction to Improve Water Quality, (2011), Achieving Nutrient and Sediment Reduction Goals in the Chesapeake Bay: An Evaluation of Program Strategies and Implementation, National Academies Press, http://www.nap.edu/catalog.php?record_id=13131.
Nebert, D., A. Whiteside and P. Vretanos, ed. (2007), OpenGIS Catalog Services Specification, OGC 07-006r1, Version 2.0.2, http://portal.opengeospatial.org/files/?artifact_id=20555.
Nemani, R., P. Votava, A. Michaelis, F. Melton and C. Milesi, (2011), "Collaborative Supercomputing for Global Change Science," EOS 92(13): 109-110, http://www.agu.org/journals/eo/eo1113/2011EO130001.pdf.
Novak, G. M. and E. T. Patterson, (1998), "Just-In-Time Teaching: Active Learner Pedagogy with WWW," IASTED International Conference on Computers and Advanced Technology in Education, Cancun, Mexico, May 27 -30, http://webphysics.iupui.edu/jitt/jitt.html.
Open Source Initiative, (2008), Open Source Initiative OSI - The BSD License, http://www.opensource.org/licenses/bsd-license.php, accessed 6/6/2010.
Piasecki, M. and B. Beran, (2009), "A semantic annotation tool for hydrologic sciences," Earth Science Informatics, 2(3): 157-168, http://dx.doi.org/10.1007/s12145-009-0031-x.
Pickett, S. T. A., M. L. Cadenasso, J. M. Grove, P. M. Groffman, L. E. Band, C. G. Boone, W. R. Burch Jr., C. S. B. Grimmond, J.Hom, J. C. Jenkins, N. L. Law, C. H. Nilon, R. V. Pouyat, K. Szlavecz, P. S. Warren and M. A. Wilson, (2008), "Beyond Urban Legends: An Emerging Framework of Urban Ecology, as Illustrated by the Baltimore Ecosystem Study," Bioscience, 58: 139-150.
Pouyat, R. V., K. T. Belt, D. E. Pataki, P. M. Groffman, J. Hom and L. E. Band, (2007), "Urban land-use change effects on biogeochemical cycles," in Terrestrial Ecosystems in a Changing World, Edited by P. Canadell, D. Pataki and L. Pitelka, The IGBP Series, Springer-Verlag, Berlin-Heidelberg-New York.
Rajasekar, A., M. Wan, R. Moore and W. Schroeder, (2006), "A prototype rule-based distributed data management system," HPDC workshop on "Next Generation Distributed Data Management", Paris, France, May 2006.
Riel, M., (1998), "Education in the 21st Century: Just-in-Time Learning or Learning Communities," Challenges of the Next Millennium: Education and Development of Human Resources. The Fourth Annual Conference of the Emirates Center for Strategic Studies and Research, Abu Dhabi, May 24-26, http://www.gse.uci.edu/vkiosk/faculty/riel/jit-learning/index.html.
SALG, (2011), Student Assessment of Their Learning Gains, http://www.salgsite.org/, accessed 7/17/11.
Schuh, P., (2004), Integrating Agile Development in the Real World, Charles River Media.
Schwaber, K., (2004), Agile Project Management with Scrum, Microsoft Press, Redmond, WA, 192 p.

Smith, M. L., W. Zhou, M. Cadenasso, M. Grove and L. E. Band, (2010), "Evaluation of the National Land Cover Database for Hydrologic Applications in Urban and Suburban Baltimore, Maryland," JAWRA Journal of the American Water Resources Association, 46(2): 429-442, http://dx.doi.org/10.1111/j.1752-1688.2009.00412.x.

Smith, P., T. Hacker and X. C. Song, (2008), "Lessons learned from Purdue campus grid infrastructure," 2nd Workshop on Desktop Grids and Volunteer Computing Systems (PCGrid 2008), Miami, Florida, April 18.

Sundaram, V., Y. Ru, L. Zhao, B. Benes, C. Song and M. Huber, (2008), "An Integrated System for Near Real-Time 3D Visualization of NEXRAD Level II Data on TeraGrid," TeraGrid 2008 Conference, Las Vegas, NV, June 9-13.

Tague, C. L. and L. E. Band, (2004), "RHESSys: Regional Hydro-Ecologic Simulation System: An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling," Earth Interactions, 8(19): 1-42, http://dx.doi.org/10.1175%2F1087-3562%282004%298%3C1%3ARRHSSO%3E2.0.CO%3B2.

Tague, C. L., L. E. Band and J. Franklin, (2006), "Terrestrial Ecosystems," in Encyclopedia of Hydrological Sciences, John Wiley & Sons, Ltd, http://dx.doi.org/10.1002/0470848944.hsa109.

Tarboton, D. G., J. S. Horsburgh, D. R. Maidment, T. Whiteaker, I. Zaslavsky, M. Piasecki, J. Goodall, D. Valentine and T. Whitenack, (2009), "Development of a Community Hydrologic Information System," 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, ed. R. S. Anderssen, R. D. Braddock and L. T. H. Newham, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009, 988-994, http://www.mssanz.org.au/modsim09/C4/tarboton_C4.pdf.

Tarboton, D. G., D. Maidment, I. Zaslavsky, D. P. Ames, J. Goodall and J. S. Horsburgh, (2010), "CUAHSI Hydrologic Information System 2010 Status Report," http://his.cuahsi.org/documents/CUAHSIHIS2010StatusReport.pdf.

Tenenbaum, D., M. Cadenasso, L. Band and S. Pickett, (2006a), "Using Transects to Sample Digital Orthophotography of Urbanizing Catchments to Provide Landscape Position Descriptions," GIScience & Remote Sensing, 43(4): 323-351, http://dx.doi.org/10.2747/1548-1603.43.4.323.

Tenenbaum, D. E., L. E. Band, S. T. Kenworthy and C. L. Tague, (2006b), "Analysis of soil moisture patterns in forested and suburban catchments in Baltimore, Maryland, using high-resolution photogrammetric and LIDAR digital elevation datasets," Hydrological Processes, 20(2): 219-240, http://dx.doi.org/10.1002/hyp.5895.

U.S. Environmental Protection Agency, (2008), Integrated Modeling Workshop 2008, Collaborative Approaches to Integrated Modeling: Better Integration for Better Decision Making, http://www.epa.gov/crem/integrated-modeling-workshop2008.html, accessed 6/5/2010.

USGS, (2007), NWIS Daily Values Web Service -- Beta 1.0, http://waterdata.usgs.gov/nwis/?DailyValues_Service_Instructions, accessed 6/13/2010.

USGS, (2010), Water Resources Applications Software, http://water.usgs.gov/software/, accessed 6/13/2010.

Vretanos, P. A., ed. (2010), Open GIS Web Feature Service 2.0 Interface Standard. OGC Implementation Standard OGC 09-025r1 and ISO/DIS 19142, Version 2.0.0, http://portal.opengeospatial.org/files/?artifact_id=39967.

Whitenack, T., D. Valentine, I. Zaslavsky, M. Piasecki, D. Tarboton, J. Horsburgh, T. Whiteaker, D. Ames and D. R. Maidment, (2010), "Hydrologic Metadata Catalog and Semantic Search Services in CUAHSI HIS ", 2010 AWRA Spring Specialty Conference Geographic Information Systems (GIS) and Water Resources VI, Orlando Florida, American Water Resources Association, Middleburg, Virginia, TPS-10-1, http://www.awra.org/orlando2010/.

Whitenack, T., I. Zaslavsky and D. Valentine, (2008), "HIS Central and the Hydrologic Metadata Catalog," Eos Trans. AGU, 89(53): Fall Meet. Suppl., Abstract IN51A-1142.

Whiteside, A. and J. D. Evans, (2008), "Web Coverage Service (WCS) Implementation Standard. OGC Implementation Standard OGC 07-067r5. Version 1.1.2.," http://portal.opengeospatial.org/files/?artifact_id=27297.

Wilkins-Diehr, N., D. Gannon, G. Klimeck, S. Oster and S. Pamidighantam, (2008), "TeraGrid Science Gateways and Their Impact on Science," Computer, 41(11): 32-41, http://dx.doi.org/10.1109/MC.2008.470.

Younts, A., A. Howard, P. Smith and J. Evans, (2009), "Bringing disruptive technology to competitions," 10th LCI International Conference on High-Performance Clustered Computing, March 10-12.
Zaslavsky, I. and D. R. Maidment, (2011), "Service orientation in the design of a community hydrologic information system," 12 in Geoinformatics: Cyberinfrastructure for the Solid Earth Sciences, Edited by G. R. Keller and C. Baru, Cambridge University Press, p.193-209.
Zaslavsky, I., D. Valentine and T. Whiteaker, (2007), "CUAHSI WaterML," OGC 07-041r1, Open Geospatial Consortium Discussion Paper, http://portal.opengeospatial.org/files/?artifact_id=21743.
Zhang, H., J. Woo, L. Zhao, D. Braun, X. C. Song and M. Lakshminarayanan, (2010), "Domain-specific web services for scientific application developers," The 6th Gateway Computing Environments Workshop (GCE10)/SC10, New Orleans, LA, November.
Zhao, L., W. Lee, X. C. Song, M. Huber and A. Goldner, (2010), "Bringing High Performance Climate Modeling into Classroom," TeraGrid conference 2010, Pittsburgh, PA, August 3-5.
Zhao, L., T. Park, R. Kalyanam, W. Lee and S. Goasguen, (2006), "The Purdue Multidisciplinary Data Management Framework Using SRB," SRB Workshop, San Diego Supercomputer Center, February 2-3.
Zhao, L., C. X. Song, V. Merwade, Y. M. Kim, R. Kalyanam, D. Ebert, B. Engel, R. Govindaraju, M. Huber, C. Jafvert, D. Niyogi and S. Prabhakar, (2007), "Interweaving Data and Computation for End-to-End Environmental Exploration on the TeraGrid," TeraGrid 2007 Conference, Madison, Wisconsin, June 4-8.
Zhao, L., K. S. K. Subramanian, C. X. Song, S. Kumar, V. Merwade, I. Chaubey, R. S. Govindaraju and M. Sayeed, (2009), "A web based interface for SWAT modeling on the TeraGrid," 5th International SWAT Conference, Boulder, CO, August 3-7.

Data Management Plan

Types of Data: Primary data collection through field or laboratory measurements is not anticipated as part of this project. However, the purpose of HydroShare is to promote collaboration and sharing of data, models, and research products, and, as such, HydroShare is necessarily designed to support the full data life cycle. We anticipate that HydroShare users will be extensively using and sharing existing data from many different sources as research products and as inputs to shared models and analyses. These data will include time series of observations from fixed monitoring locations (e.g., streamflow gages, weather stations, etc.), geospatial datasets (e.g., land use, digital elevation models, etc.), and potentially time varying gridded data such as radar-based precipitation products.

To facilitate shared access to the full range of anticipated data types and sources, significant development efforts on this project will be devoted to creating linkages between HydroShare and data providers such as NASA and the USGS National Water Information system. HydroShare must also facilitate sharing of individual investigator datasets and derived data products within a collaborative environment. Development of the required repositories, protocols, workflows, and methods for doing so are part of the intellectual contribution of this project. Where possible, we will use existing web services (e.g., those available already through the CUAHSI HIS) or develop data services and workflows that interface with existing, Internet-based systems for retrieving the primary data that users will require (e.g., the USGS Seamless Data Server for the National Land Cover or National Elevation Datasets), increasing the availability of these datasets to data consumers working within HydroShare.

It is also anticipated that HydroShare users will generate large quantities of model simulation results. We anticipate that researchers will publish model “packages” consisting of the model code, model inputs (e.g., the set of model input data that can be used to execute the model for a particular scenario), and model simulation results to ensure that their work is transparent and reproducible. We anticipate that hydrologic model simulation results will be of types similar to those described above. Within HydroShare, we plan to enable the sharing of not only data and model codes, but also other collaborative products such as codes for analytical tools and research results that may consist of images, documents, and other digital materials resulting from group collaborations.

Data and Metadata Standards: Data services in HydroShare will make full use of existing and emerging standards for sharing environmental datasets. For example, we will use CUAHSI HIS WaterOneFlow web services for time series of point observations and will transmit data in Water Markup Language (WaterML) format (Zaslavsky et al., 2007). For geospatial datasets, we will use existing Open Geospatial Consortium (OGC) standard interfaces such as Web Map Services (de la Beaujardiere, 2006), Web Feature Services (Vretanos, 2010), and Web Coverage Services (Whiteside and Evans, 2008). Metadata cataloging services will use existing CUAHSI HIS services and OGC Cataloging Services for the Web (Nebert et al., 2007). HydroShare will also provide standard HTTP and/or FTP access to download datasets and research products that are more easily used as files rather than services, but will ensure that published products are accompanied by appropriate metadata descriptions either delivered by the data service or accompanying the download. Metadata descriptions for all datasets will conform to appropriate ISO and INSPIRE metadata specifications with extensions for the detailed provenance information for model results.

An important aspect of our project will be our work to advance metadata standards for the description of contributed models, tools, and other research products. Here we will use our experience with developing standards via the CUAHSI HIS to specify the structure and contents of model metadata packages in such a way that contributed models and tools can be published for sharing among collaborative groups and potential model users.

Policies for Data and Research Products: The goal of HydroShare is to promote collaboration and sharing of data, models, and analyses though the NSF supported environments of CUAHSI HIS, HUBzero, and iRODS. Groups of researchers may wish to share data, model input packages, or simulation results within their group before they are published externally. HydroShare will provide users with the choice to create public or private collaboration spaces accessible only to selected users within which these activities can take place. Authentication and access control will be fully integrated within HydroShare. We expect that collaborative activities may result in large quantities of data, only a fraction of which may be considered publishable by the researchers. As such, HydroShare will establish the facilities required to publish data, models, and simulation results, enabling individual researchers to select and publish their results as they see fit. Access to the contents of private research group working areas will be at the discretion of group members to protect the intellectual property that they create. Final research results can be made freely and publicly available when they are deemed publication ready by the author, and all products published in HydroShare will include appropriate attribute and citation information. Published products may also include an access/use agreement specified by the author that is in conformance with established practices and policies already in place among the collaborating Universities and CUAHSI.

As a general policy, all source code developed by this project will be created using an open development model and will be distributed under the New BSD (BSD 2) Open Source license. We will use open-source code repositories for our software development, including an instance of gForge hosted by RENCI. This will enable us to coordinate our development activities across multiple Universities and engage developers and contributors from outside of the immediate project team who wish to contribute.

Plans for Archiving Data: As stated above, researchers will curate and archive data, derived data products, models and tools, and model input and output data packages as research products within HydroShare to better enable collaboration and reproducibility of research results. Two mechanisms will enable this. First, HydroShare itself will serve as the primary archival mechanism for curated products created by community researchers. HydroShare and all of its attendant systems will be hosted on fault-tolerant, enterprise-class servers dedicated to this project and housed in RENCI's managed, climate controlled, UPS-backed IT facility ensuring the reliability of the HydroShare system.

HydroShare will contain a hydrology body of knowledge that will be valuable to other research communities and discoverable and exposed via published spatial information standards. As an additional archival mechanism, we will establish within HydroShare appropriate interfaces that enable curated research products archived in HydroShare to be indexed by the cyberinfrastructure being built by the NSF-funded DataNet projects. For example, we will establish the necessary web service interfaces for HydroShare to become a DataONE Member Node (DataONE Project Team, 2011). By creating a DataONE Member Node, HydroShare can participate in the robust, distributed DataONE network and take advantage of the indexing, archival, and discovery services that DataONE provides, broadening the impact of published HydroShare resources and encouraging their reuse. Similarly, federation with the DataNet Federation Consortium (DFC) that has recently been recommended for funding can be readily accomplished due to the reliance of HydroShare and DFC on iRods. We anticipate that curated research products published in HydroShare will be citable for use in peer-reviewed journal articles and conference presentations and proceedings.

Postdoctoral Researcher Mentoring Plan

One postdoctoral fellow hired for this project at the University of North Carolina will be integral and centrally involved in this research. The project will provide unique mentoring and training experiences for the fellow spanning the area of advanced informatics, simulation modeling and environmental science. The fellow will be resident in the Institute for the Environment at UNC, but will work closely with project personnel at RENCI, as well as at partner universities. Co-PI Band will take primary responsibility for advising the fellow on work scheduling, career goals, and developing networks and opportunities for successful career development. Weekly meetings with Band's research group, and with RENCI personnel will be augmented by quarterly web or in-person meetings with the full project team.

The fellow will work collaboratively with the Baltimore Ecosystem Study and Coweeta LTER sites (co-funded by both), maximizing exposure and experience with existing NSF funded observatories, while also receiving training and experience in the development and use of advanced cyberinfrastructure tools. As a result, the fellow will gain first-hand experience with large, organized research facilities, academic institutes, and the interface between environmental science and software development. The fellow will integrate research expertise in the development and use of sustainable software supporting ecohydrological modeling and analysis tools as their major project roles.

We envision two post-doctoral fellows serving in this position over the five year term of the project, with 2-3 year periods. Both fellows will contribute to development and application of the project cyberinfrastructure tools, although the second fellow will be primarily responsible for translating these tools into fundamental new science. The synthetic training gained in the project will produce opportunities for employment in the university, government agency or private industry sectors.

Management and Coordination Plan - Collaborative Research: SI2-SSI: An interactive software infrastructure for sustaining collaborative community innovation in the hydrologic sciences

Management and Coordination Plan

The purpose of this Management and Coordination Plan is to describe the mechanisms for integrating the proposed research across institutions and for coordinating and executing project activities. Methods, infrastructure and budget references are discussed. The specific roles of the PI, co-PIs, senior personnel and other personnel at all institutions involved are detailed below.

Roles and Responsibilities

Dr. David Tarboton of Utah State University (USU) is the overall PI and **Project Director** and USU is the lead institution. Dr. Jennifer Arrigo, CUAHSI Program Manager, serves as co-PI on the project. Subcontracting to USU are Idaho State University, University of Texas at Austin, Tufts University, CUAHSI, and University of South Carolina. Additionally, Larry Band and a post-doc from University of North Carolina (UNC) Chapel Hill will be part of this team but subcontract through UNC. The individuals at these institutions collectively represent the hydrology domain science, user interface design and usability testing, and research aspect of the proposal.

RENCI at the University of North Carolina (UNC) Chapel Hill is submitting as a NSF Collaborating Research institution with USU and Ray Idaszak is PI of this effort and an overall proposal co-PI. Subcontracting to UNC Chapel Hill are the San Diego Supercomputer Center (SDSC) at the University of California San Diego (UCSD), and Purdue University. The individuals at these institutions collectively represent the cyberinfrastructure and computer science aspect of the proposal.

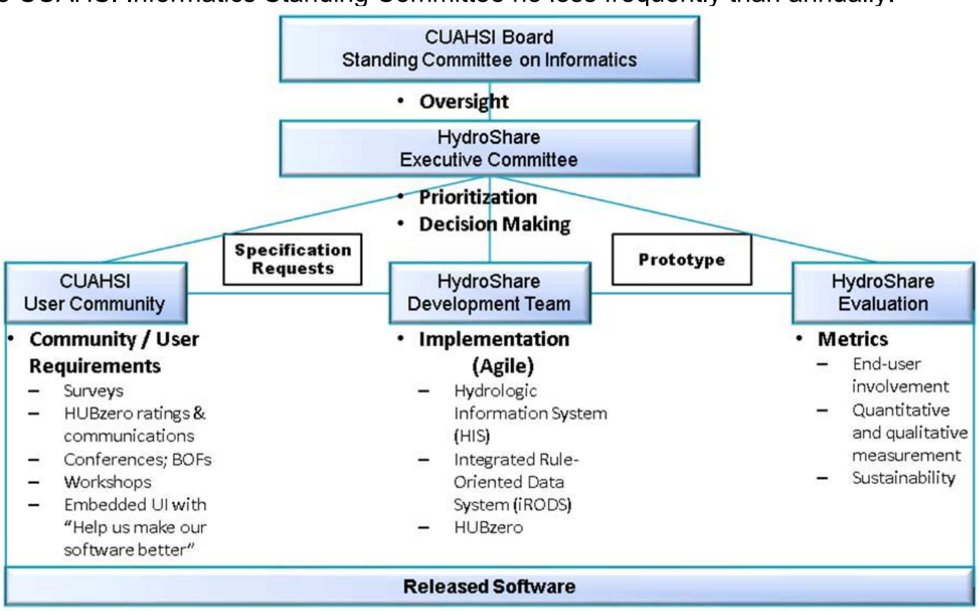
The specific roles and responsibilities are summarized in the following table.

Participant	Roles and Responsibility
Hydrology Domain	
Utah State University (David Tarboton, Jeff Horsburgh, Kim Schreuders)	Overall project direction. User interface. Data interface and format specification (decisions as to what types of data to support). Data model. Digital hydrology framework. Interface with DataOne and establish DataOne member node (Horsburgh).
Idaho State University (Dan Ames)	HydroDesktop extensions to interact with HydroShare. Create and publish data packets. Access online data using web services. Configure and publish model workflows. HUBzero domain-specific use cases.
University of Texas, Center for Research in Water Resources (David Maidment, Tim Whiteaker)	Coordination with agencies and partners. Prototyping and use testing. Development of river cross section exemplar.
Tufts University (Alva Couch)	User Interface design. User Interface usability testing. Search and tagging.
CUAHSI (Jennifer Arrigo, Yoori Choi)	Community support, education, training, and outreach.
University of South Carolina (Jon Goodall)	Model metadata representation. Define model description semantics. Liaise to CSDMS.
UNC Chapel Hill – Institute for the Environment (Larry Band, postdoctoral fellow)	Modeling with RHESSys. Integrate RHESSys components as a demonstration and use to leverage existing NSF Observatory science. Liaise to NASA Earth Exchange.
Cyberinfrastructure	
UNC Chapel Hill - RENCI (Ray Idaszak, Jason Coposky, Jason Reilly, Dan Bedard)	IT Coordination. Overall software integration, engineering and interfaces between the various technologies (iRODS, HUBzero, CUAHSI HIS, Web Services, adherence to standards). Integrate RENCI OSG EngageVO, workload

	management, grid workflow to HydroDesktop and HUBzero. Collaborative Development Environment. Build and test environment. Liaise to Open Science Grid, TeraGrid / XSEDE, and DataNet Federation Consortium.
UCSD SDSC (David Valentine, Tom Whitenack)	Web services and HydroCatalog integration with iRODS. Further development of OGC-compliant water data services and implementation in HydroShare.
Purdue (Venkatesh Merwade, Carol Song)	HubZero development based on capabilities in WaterHUB. Add social-net capabilities of rating published data and models. Add social-net capabilities of sharing collaborative comments. Add "grouping" capabilities so groups can use HydroShare to collaborate around models. Link to HydroDesktop such that it is easy and straightforward to publish models in HydroDesktop to HydroShare.

Project Management Across Institutions and Disciplines

The project governance is depicted in the figure below. CUAHSI represents the hydrology community and has a Standing Committee on Informatics appointed by the CUAHSI Board. This committee oversees Informatics projects such as this. This committee is chaired by Paul Houser (membership is listed at http://www.cuahsi.org). To gain community input on the direction and priorities of the project we will submit the project for regular review by the Informatics Standing Committee of CUAHSI who will effectively serve as an Advisory Committee for the project. CUAHSI is also setting up a HIS Users Committee that will provide more detailed feedback on specific functionality and usage testing. To govern this project we will create a HydroShare Executive Committee comprised of Tarboton, Arrigo, and Idaszak representing the domain science, community and CI respectively. The HydroShare Executive Committee will meet (in person, by telephone or videoconference) at least once a month. The HydroShare project will report to the CUAHSI Informatics Standing Committee no less frequently than annually.



User requirements will be gathered via the mechanisms shown in the diagram. The HydroShare Evaluation Metrics described in proposal Section 8 will be used to evaluate the degree to which every release of the HydroShare Software is meeting the expectations of the CUAHSI user community. Development and support of Released Software is discussed in the next section.

Coordination Mechanisms

The hydrology team members have a long history of collaboration, communication, and interactions that will benefit the project. All PI's interact frequently on a variety of projects and working groups.

Tarboton will serve as the overall Project Director ensuring that project deliverables are met. Tarboton will organize yearly PI meetings utilizing proposal travel budget. Tarboton will also organize bi-weekly project conference calls. Recurring CUAHSI workshops such as the CUAHSI Conference on Hydrologic Data and Information Systems and CUAHSI Biennial Colloquium on Hydrologic Science and Engineering will be used to present HydroShare activities and progress, and organize project participants, partners, and users as an opportunity to understand and gather user requirements.

Idaszak will serve as Project Manager for the cyberinfrastructure team including the software engineering process and testing methodology as described in Section 5 of the proposal. RENCI, the lead cyberinfrastructure institution has extensive institutional experience in software development and coordination and management of large projects. The software engineering process we will use addresses changing technology and feature requirements in a manner resulting in long-term sustainability.

We will use the Agile software development methodology (Schuh, 2004) and Scrum framework (Schwaber, 2004; Gorakavi, 2009) to enable our geographically distributed team to collaborate in developing HydroShare. A key principle of the Agile method and Scrum framework is recognizing that the complete product functional requirements cannot be fully anticipated, hence the team must be organized and software engineering tools emplaced to explicitly support evolutionary and adaptive identification of needs, implementation of solutions, deployment to end users, and assessment through feedback.

These practices have been in place and actively used by the iRODS@RENCI team that co-PI Idaszak oversees in the development of an enterprise version of iRODS (Brieger et al., 2011). The Agile-Scrum approach has three phases including pre-game, development, and post-game.

The Pre-game phase includes planning, architecture, high level design, and the creation of a product backlog list. This list includes all the requirements as user “stories.” User stories originate from stakeholders in coordination with domain and technology experts. Since the team consists of geographically distributed groups, we will use tools that support collaborative maintenance of product backlogs. The members of the HydroShare Executive Committee will be involved in the pre-game phase decisions and kept apprised of sprint activities, discussed below, through a Scrum dashboard. In the development phase, HydroShare will be developed in sprints. The architecture and design of the system evolve during “sprints” which consist of iterations over a set number of weeks. A working build will be available at the end of every sprint for demonstration and evaluation as part of the description of the sprint. A specified number of sprint cycles in one system development process results in a system ready for distribution and testing. Each sprint includes the traditional phases of software development: requirements, analysis, design, development and testing, and delivery. The post-game phase puts closure to a release. This phase is entered when all requirements for a given release are completed and have passed system and integration testing. A code snapshot is captured and tagged. Product documentation, based on user stories and in-source documentation written in the development phase, is completed here. Training of product support and maintenance teams is performed in this phase. This is critical for sustainable product implementation. Training will be given to quality assurance (QA) and outreach team members on the new functionality developed, for incorporation into education and outreach activities detailed below. Feedback from this training, both from outreach team members and the community, will provide critical input for future product development.

Standard source code management practices including code versioning, branching and tagging will be used in a continuous integration process. We will use industry standard source code management tools such as those included in SourceForge, GitHub, or CodePlex. Currently the CodePlex.com open source collaboration portal is used for HydroDesktop and HydroServer.

Two HydroShare sites, a production site and a beta or development site will be maintained. The beta-site will be used to identify bugs, perform verification and validation, conduct performance tests against defined metrics, and survey overall usability prior to making each release available on the production site.

Budget References

The HUBzero hardware acquisition to host HydroShare is covered in Section 3 of the UNC Chapel Hill budget justification. The proposed Domestic travel budget will be used for participation in the annual NSF PI meeting, and also for team members to attend hydrology conferences and workshops to support the training, education and outreach for the proposed system.