Design and Development of a Prototype Client-Server
Water Data Sharing Software System

Cuyler S. Frisby

A project submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Daniel P. Ames, Chair
Norman L. Jones
E. James Nelson

Department of Civil and Environmental Engineering
Brigham Young University
June 2014

Copyright © 2014 Cuyler S. Frisby
All Rights Reserved
ABSTRACT

A HydroShare Plugin for HydroDesktop:
Data Sharing Made Easy

Cuyler S. Frisby
Department of Civil and Environmental Engineering, BYU
Master of Science

The new NSF-funded HydroShare project has as a primary goal the dynamic and interactive sharing of hydrologic and climate data and models - primarily through a web-based end user interface. In addition to the HydroShare web interface, we have proposed adding a number of capabilities to the CUAHSI Hydrologic Information System (HIS) HydroDesktop software tool (www.hydrodesktop.org) to support interaction with HydroShare directly from the user desktop. Our goals for extending HydroDesktop as a client for HydroShare include: 1) continuing to support key HIS data functions: search, download, view, export; 2) including new support for creating “data packages” from observed data and model output and uploading these data packages to HydroShare; 3) searching HydroShare for data packages which can be loaded directly into HydroDesktop; and 4) tracking changes to data and models, maintaining this provenance information, and posting it to HydroShare. This paper presents our approach to adding this new functionality to HydroDesktop including creating a relatively light-weight plugin which exposes a new ribbon user control and buttons and using Open Geospatial Consortium (OGC) web services standards, as well as custom web services to retrieve and post data.

Keywords: GIS; desktop applications; data sharing, hydroinformatics, CUAHSI, HydroShare, HydroDesktop
TABLE OF CONTENTS

1 Introduction.................................................................................................................. 1
   1.1 Problem Statement.................................................................................................... 1
   1.2 Background ............................................................................................................. 2
   1.3 Research Goals ...................................................................................................... 10
2 Prototype Design and Development ........................................................................... 13
   2.1 Use Cases for HydroShare Ribbon for HydroDesktop ........................................... 13
   2.2 Selecting a Programming Language ....................................................................... 18
   2.3 Development of a Prototype HydroShare Ribbon for HydroDesktop .................... 21
   2.4 Contributions to the HydroShare Website Development ........................................ 23
3 Completed Research Paper – Client-Server Hydrologic Data Exchange – Making the
   (Use) Case for Desktop Applications ............................................................................ 25
   3.1 BACKGROUND ...................................................................................................... 25
   3.2 METHODS ............................................................................................................. 28
      3.2.1 Defining a “HydroDesktop Project” resource type ........................................... 28
   3.3 RESULTS ............................................................................................................... 32
      3.3.1 Server-Side Design (HydroServer Website) ....................................................... 32
      3.3.2 Client-side design (HydroDesktop) ................................................................... 33
   3.4 CONCLUSION ........................................................................................................ 37
   3.5 iEMSs Conference Review Process ....................................................................... 37
4 Conclusions .................................................................................................................. 41
REFERENCES .................................................................................................................. 43
LIST OF FIGURES

Figure 1. CUAHSI HIS Infrastructure .................................................................2
Figure 2. HydroDesktop User Interface Showing the Provo, Utah Area ...............4
Figure 3. HydroDesktop Search Tab ..................................................................5
Figure 4. Locations of Streamflow Time Series Matching the Search Parameters ....6
Figure 5. Viewing of Tabular Data in HydroDesktop ..........................................7
Figure 6. Viewing of Graphical Data in HydroDesktop .......................................8
Figure 7. Resource Repository Centric Approach for Modeling and Analysis ........9
Figure 8. Facilitating the Interaction Between HydroDesktop and Python via IronPython ..20
Figure 9. Original Prototype HydroShare Ribbon in HydroDesktop .......................20
Figure 10. Launching a Python Script from HydroDesktop via IronPython ............21
Figure 11. Prototype HydroShare Ribbon in HydroDesktop ....................................22
Figure 12. Prototype Python-Based HydroShare Download Tool in HydroDesktop ....22
Figure 13. HydroShare Web Service URLs in Python Downloader Source Code ........23
Figure 14. Files Associated with a Sample HydroDesktop Project .........................29
Figure 15. HydroDesktop Project in a BagIt File ..................................................32
Figure 16. HydroShare Ribbon in HydroDesktop ..................................................33
Figure 17. HydroShare Download Form in HydroDesktop .....................................34
Figure 18. HydroShare Login Screen in HydroDesktop .........................................35
Figure 19. Resource Type Selector Form in HydroDesktop ....................................35
Figure 20. Metadata Entry Form in HydroDesktop ..............................................36
1 INTRODUCTION

1.1 Problem Statement

No natural resource is more critical to life than water, which interacts with many components of the Earth, such as ecosystems, climate, geologic processes, economics, and human health in complex, connected ways. “A central challenge in the hydrologic sciences is scaling measurements and models from the level of the lab or local field to site to regional and global extents” (CUAHSI 2010). “As scientists begin to investigate complex hydrologic processes at expanding spatial and temporal scales, integration of data from multiple sources, projects, and research efforts becomes critical” (Ames, Horsburgh et al. 2012). The amount of water-related data collected across the globe continues to increase. However, the hydrologic sciences are severely hindered by a lack of efficient tools for the sharing and discovery of hydrologic resources and data. (CUAHSI 2010).

New scientific discoveries are often catalyzed by applying new technology and data resources to existing problems. By aggregating results from local research projects across sites and times, the potential exists to advance science and research significantly through the publication of research data. There is a need for standardized and robust methods to publish and organize hydrologic data in a way that can be discovered and used for scientific analysis (Horsburgh, Tarboton, et al. 2009).
1.2 Background

The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) is a 501(c)3 research organization funded by the National Science Foundation (NSF) representing more than 100 U.S. universities and international water science-related organizations. CUAHSI has as its main goal the advancement of water science in the United States. This includes improving access to hydrologic data, information and models and facilitating interactions among the diverse water research community (CUAHSI, 2010). To this end, the CUAHSI Hydrologic Information System (HIS) is an internet-based system for sharing hydrologic data. It is comprised of databases and servers, connected through web services, to client applications, allowing for the publication, discovery and access of data (Horsburgh, Tarboton, et al. 2009). Figure 1 below shows the main components of the CUAHSI HIS infrastructure and their relationship to one another.

![CUAHSI HIS Infrastructure](image)

Figure 1. CUAHSI HIS Infrastructure (“CUAHSI Hydrologic Information System”, 2008-2010)

As shown above in Figure 1, there are three key hardware/software components that store and process data within the CUAHSI HIS infrastructure:
1. HIS Central – contains copies of metadata which facilitates searches; works like a search engine, in that it harvests metadata from the data servers and allows it to be efficiently searched by the clients.

2. HydroServer – stores, organizes, and publishes data; allows metadata to be harvested by HIS Central and data to be shared with clients.

3. Clients (such as HydroDesktop) – gives users a convenient interface to access data; retrieves metadata form HIS Central and retrieves data from HydroServers.

There are also three types of web services that allow computers within the HIS to communicate via the internet:

1. Data Services – allows water and related spatial data to be retrieved by client (such as HydroDesktop) computers.

2. Search Services – allows client computers to perform searches of the search catalog at HIS Central.

3. Metadata Services – allows HIS Central to retrieve the metadata necessary to build the search catalog (“CUAHSI Hydrologic Information System”, 2008-2010).

The primary client application of the HIS infrastructure is HydroDesktop, which is a free and open source desktop application which gives users a convenient interface to access data, retrieves metadata from HIS Central and downloads data from HydroServers (Ames, Horsburgh, et al. 2012).

“When HydroDesktop opens, the user is presented with a map showing political boundaries included as a GIS dataset in HydroDesktop. The user can augment this map by displaying a variety of online basemaps from ESRI, Google, Bing, OpenStreetMap, and more. By manipulating and
navigating the map, the user sets the spatial context for the study,” (Ames, Horsburgh, et al. 2012) as shown below in Figure 2.

Figure 2. HydroDesktop User Interface Showing the Provo, Utah Area

Once zoomed in to the area of interest, “the user activates the Search tab to specify filters for the data search. In this case, the user draws a box (shown in red in Figure 3 below) around the area of interest, specifies a keyword such as “streamflow,” and enters a date range in order to constrain the search. The user could further filter the search to only include specific data sources of interest, but in this case will leave the default of search across all … web services registered at HIS Central” (Ames, Horsburgh, et al. 2012).
“With the search parameters set, the user clicks Run Search, which triggers a query to HIS Central for metadata about time series that match the search criteria. HydroDesktop presents the search results to the user as points on the map where each point is symbolized by data source. Larger points indicate a greater number of time series values available at that location compared to locations with smaller points (See Figure 4 Below). The user can hover the mouse over a point to view metadata about the time series available at the point or open a table to view metadata about the time series at all locations” (Ames, Horsburgh, et al. 2012).
The user can then select a data site and click the “download” button to download the time series. “HydroDesktop identifies the URL to data source from the metadata and makes a connection to that data source’s … web service to download the data, saving the result to the local database” (Ames, Horsburgh, et al. 2012).

With the data retrieved, the user can then click the “Table” tab to view the data in tabular form as shown below in Figure 5.
The user can also click the “Graph” tab to view the data in graphical form, as shown below in Figure 6. The user can select from several types of plots: time series, probability, histogram, and box/whisker.
HydroDesktop also supports extensions or “plug-ins” which allow for more sophisticated analysis within the HydroDesktop environment. “HydroDesktop includes a number of plug-ins developed by the HIS team, and also supports third party plug-ins that follow a standard, well-defined plug-in interface described at the project website” (Ames, Horsburgh, et al. 2012). Such a plug-in will be built to allow interaction between HydroDesktop and a new website being developed, HydroShare.

HydroShare is an NSF-funded research project which seeks to expand the data sharing capability of the CUAHSI HIS “by broadening the classes of data accommodated, creating new capability to share models and model components, and taking advantage of emerging social media functionality to enhance information about and collaboration around hydrologic data and models” (Tarboton, Idaszak et al. 2014). The new website HydroShare.org, being created in conjunction
with CUAHSI, will be “an online, collaborative system” (Tarboton, Idaszak et al. 2014) which will “enable scientists to easily discover and access hydrologic data and models, retrieve them to their desktop or perform analyses in a distributed computing environment that may include grid, cloud, or high performance computing instances as necessary. Scientists may also publish outcomes (data, results or models) into HydroShare, using the system as a collaboration platform for sharing data, models and analyses” (Tarboton, Idaszak et al. 2014).

Each discrete unit of data on HydroShare will be represented as HydroShare Resource, a fundamental concept in HydroShare which is represented by a Resource Data Model (Tarboton, Idaszak et al. 2014). A resource “can be data sets in any of the supported formats, models, or just generic resources comprising a file or files not useable to the system, but important to a user’s project” (Tarboton, Idaszak et al. 2014). HydroShare is being built using a “resource centric approach,” in which “tools or models perform actions (analyses) on resources in the resource store that are in a standard format” (Tarboton, Idaszak et al. 2014). This approach is shown below in Figure 7.

Figure 7. Resource Repository Centric Approach for Modeling and Analysis (Tarboton, Idaszak et al. 2014)

HydroShare seeks to provide an easy-to-use web-based interface for the sharing and discovery of hydrologic data. HydroShare users will be able to upload resources to the website
which will then be discoverable by other users. This new paradigm for the sharing of hydrologic data will allow individuals with very different research objectives from around the world to collaborate and support one another (Tarboton, Idaszak, 2012).

1.3 Research Goals

HydroDesktop currently lacks a means whereby to interact with the new HydroShare website. The goal of my research is the design and development of a prototype HydroShare plugin which will allow the discovery and sharing of resources from within the HydroDesktop environment. Users will be able to search for shapefiles, data series, and projects and download these directly into their current HydroDesktop project. They will also be able to upload their current project or an individual data layer to the HydroShare website. The development of this plugin will be of great interest to users of HydroDesktop as well as hydrologic professionals in general. It will greatly simplify the sharing and discovery of data, increasing the pace at which hydrologic science can progress.

The general goals for a HydroShare plugin in HydroDesktop fit under the primary goals of CUAHSI. They include “improving access to hydrologic data, information and models and facilitating interactions among the diverse water research community” (CUAHSI, 2010). Thus, the first goal for the HydroShare plugin is to provide for the discovery of hydrologic resources located on HydroShare.org. HydroDesktop already provides for the discovery and viewing of hydrologic data. This, in fact, is the primary objective of HydroDesktop (Ames, Horsburgh, et al. 2012). However, this new HydroShare plugin will extend that functionality to include resources located on HydroShare, whereas it currently only supports resources registered with HIS Central (Ames, Horsburgh, et al. 2012). The plugin will allow users to search for resources
using filters such as file type and keyword search. It will also allow for the download of these resources. These resources can be imported into the current HydroDesktop project immediately after download.

The other primary function of the HydroShare plugin will be uploading of resources. This may include time series data, data layers, scripts, and entire HydroDesktop projects. The ability to upload these resources directly from the HydroDesktop interface will greatly simplify the process of sharing data. As a result, it is expected that the volume of shared data will greatly increase through the implementation of this plugin. This will meet the primary goals of CUAHSI and will increase the collaboration and support among hydrologic professionals (CUAHSI 2010).

To facilitate the seamless download and upload of data using the HydroShare plugin, several technical issues must be addressed. Firstly, users are required to create accounts HydroShare before they can access data or upload their own resources. Thus, user authentication must be built into the HydroShare plugin. Users will be prompted for their HydroShare username and password when they attempt to access HydroShare, either for download or upload purposes. This method provides the added benefit of metadata fields for the uploaded resource being automatically filled from the information associated with the user’s account.

A method to package all of the data for a HydroDesktop project will also be needed. It is hoped that individual base layers can be stored on HydroShare and shared among users. For example, if a user includes a “U.S. States” layer in their project, when they upload their project it will simply point to the copy of the layer located on HydroShare rather than uploading another copy. This reduces the redundancy of information stored on the HydroShare servers and increases the efficiency of the system.
The remainder of this project report is organized as follows. Chapter 2 presents the design and development of the prototype HydroShare plugin for HydroDesktop. Chapter 3 presents a complete technical paper prepared for presentation at the 2014 American Water Resources Association Spring Specialty Conference GIS and Water Resources, Snowbird, Utah (oral presentation) and the 7th International Environmental Modeling and Software Congress in San Diego, California (poster presentation). Chapter 4 summarizes the key contributions of the work and presents opportunities for future development.
2 PROTOTYPE DESIGN AND DEVELOPMENT

As a prototype design and development project, this effort required a significant research effort as well as a number of “trial and error” components. For example, the very programming language to be used in the project was in question from the beginning of the effort. Additionally, it was critical to develop a series of “use cases” describing how users will use the tools. A working prototype of the HydroShare plugin for HydroDesktop was developed with a few working functions, however the server component (the HydroShare web site) was completely redesigned in January 2014 and all previously working functionality of the HydroDesktop plugin became unusable. A summary of these activities is given in this chapter.

2.1 Use Cases for HydroShare Ribbon for HydroDesktop

The following use cases were written to demonstrate the steps a user of HydroDesktop would go through when using the HydroShare ribbon. A use case is a written description of how users will perform tasks on a website. Each use case is presented as a series of simple steps, beginning with the user’s goal and ending when that goal is fulfilled. Use cases add value because they explain how the system should behave. They also help brainstorm what could potentially go wrong with the system (“Use Cases” n.d.). These use cases were presented to the HydroShare research team during one of the weekly telephone conference calls. These
documents encompass the primary functions which will be implemented in the HydroShare ribbon for HydroDesktop.

<table>
<thead>
<tr>
<th>ID:</th>
<th>Title:</th>
<th>User Uploads a Data Layer to HydroShare from within HydroDesktop</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Description:</td>
<td>A HydroDesktop user wishes to upload an individual HydroDesktop data layer to HydroShare. The user navigates to the HydroShare ribbon in HydroDesktop and logs in to HydroShare. The user uploads the resource to HydroShare.</td>
</tr>
<tr>
<td></td>
<td>Stakeholders:</td>
<td>The user that wishes to upload a data layer to HydroShare Any HydroShare users or organizations which may download and use the HydroDesktop data layer after it is uploaded to the HydroShare website.</td>
</tr>
<tr>
<td></td>
<td>Actors:</td>
<td>The user that wishes to upload a data layer to HydroShare The HydroShare system manages access to HydroShare resources.</td>
</tr>
<tr>
<td></td>
<td>Trigger:</td>
<td>A user clicks the “Upload” button on the HydroShare ribbon in HydroDesktop</td>
</tr>
<tr>
<td></td>
<td>Preconditions:</td>
<td>The user is logged into HydroShare The user is authorized by the HydroShare system to upload resources</td>
</tr>
<tr>
<td></td>
<td>Postconditions:</td>
<td>The HydroDesktop data layers are uploaded to HydroShare and can be discovered and downloaded by other HydroShare users</td>
</tr>
<tr>
<td></td>
<td>Main Success Scenario:</td>
<td>A user logs into the HydroShare website via the HydroShare ribbon in HydroDesktop The user clicks the “upload” button located on the HydroShare ribbon in HydroDesktop From a list of options, the user specifies that a specific data layer will be uploaded to the HydroShare website All relevant files associated with the HydroDesktop data layer are zipped and uploaded to the HydroShare website</td>
</tr>
<tr>
<td></td>
<td>Extensions:</td>
<td>If the user is not logged in, HydroDesktop prompts the user to log in to HydroShare before uploading any HydroShare resources If the user does not have a HydroShare account, a link is provided to create one.</td>
</tr>
<tr>
<td></td>
<td>Approved for Implementation:</td>
<td>[Yes or no. Who approved. When approved.]</td>
</tr>
<tr>
<td></td>
<td>Change Log:</td>
<td>Cuyler Frisby 1/21/2014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ID:</th>
<th>Title:</th>
<th>User Downloads a HydroDesktop Project from HydroShare via the HydroShare Website</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14
### Description:
A HydroShare user wishes to download a complete HydroDesktop project from HydroShare. The user navigates to the HydroShare website and logs in to HydroShare. The user searches for and downloads the desired project from HydroShare.

### Stakeholders:
The user that wishes to download a HydroDesktop project from HydroShare. The HydroShare user or organization who uploaded the HydroDesktop project to the HydroShare website.

### Actors:
The user that wishes to download a HydroDesktop project from HydroShare. The HydroShare system manages access to HydroShare resources.

### Trigger:
A user clicks the “Download” button on the information page for a HydroDesktop project on HydroShare.

### Preconditions:
The user is logged into HydroShare. The user is authorized by the HydroShare system to download resources.

### Postconditions:
The HydroDesktop project files are downloaded to a user-specified location on a local computer.

### Main Success Scenario:
- A user logs into the HydroShare website.
- The user searches for the desired HydroDesktop project using the “Find Resources” section of the website.
- A list of available resources which meet the search criteria is returned.
- The user selects a HydroDesktop project from the list and is presented with information regarding the project.
- The user downloads the HydroDesktop project by clicking the “download” button.
- The HydroDesktop project files are downloaded and saved to a specified location on the user’s computer.

### Extensions:
If the user is not logged in, HydroDesktop prompts the user to log in to HydroShare before searching for HydroDesktop projects. If the user does not have a HydroShare account, a link is provided to create one. A preview of the HydroDesktop project is shown on the HydroShare website, including a screenshot of the project.

### Change Log:
Cuyler Frisby 1/21/2014
for and downloads the desired project from HydroShare and the project appears in the HydroDesktop project window.

**Stakeholders:**
The user that wishes to download a HydroDesktop project form HydroShare
The HydroShare user or organization that uploaded the HydroDesktop project to the HydroShare website.

**Actors:**
The user that wishes to download a HydroDesktop project from HydroShare.
The HydroShare system manages access to HydroShare resources.
The HydroDesktop application allows users to search for and download HydroDesktop projects from HydroShare.

**Trigger:**
A user clicks the “Download Resource from HydroShare” button on the HydroShare ribbon of HydroDesktop.

**Preconditions:**
HydroDesktop is running
The user is logged into HydroShare (optional for downloading)
The user is authorized by the HydroShare system to download resources

**Postconditions:**
The new HydroDesktop project appears in the main window of HydroDesktop.

**Main Success Scenario:**
A user logs into HydroShare through the HydroShare ribbon in HydroDesktop
The user clicks on the “Download from HydroShare” button on the HydroShare ribbon in HydroDesktop and is presented with a search box and several filter options such as “resource type” and “keyword”
The user executes a search query and a list of available HydroShare resources which meet the specified criteria is returned
The user selects the desired HydroDesktop project and clicks the “download” button
The HydroDesktop project is downloaded and appears in the main window of HydroDesktop

**Extensions:**
If the user is not logged in, HydroDesktop prompts the user to log in to HydroShare before searching for HydroDesktop projects.
If the user does not have a HydroShare account, a link is provided to create one.
If a HydroDesktop project was already open, HydroDesktop prompts the user to save it before a new project is downloaded.

**Approved for Implementation:**
[Yes or no. Who approved. When approved.]

**Change Log:**
Cuyler Frisby 1/21/2014

| ID: | Title: | User Downloads a Data Layer from HydroShare via HydroDesktop |
Description: A HydroDesktop user wishes to add a data layer to their project. From within HydroDesktop, the user searches for and downloads the desired files from HydroShare and they are added automatically to the current project.

Stakeholders: The user that wishes to download a data layer from HydroShare. The user or organization that uploaded the data layer to the HydroShare website.

Actors: The user that wishes to download a data layer from HydroShare. The HydroShare system manages access to HydroShare resources. The HydroDesktop application allows users to search for HydroShare resources and add them to the current project.

Trigger: A user clicks the “Download Resource from HydroShare” button on the HydroShare ribbon of HydroDesktop.

Preconditions: HydroDesktop is running. The user is logged into HydroShare (optional for downloading). The user is authorized by the HydroShare system to download resources.

Postconditions: The new data layer appears as a layer in the current HydroDesktop project.

Main Success Scenario: A user logs into HydroShare through the HydroShare ribbon in HydroDesktop. The user clicks on the “Download from HydroShare” button on the HydroShare ribbon in HydroDesktop and is presented with a search box and several filter options such as “resource type” and “keyword.” The user executes a search query and a list of available HydroShare resources which meet the specified criteria is returned. The user selects the desired resource and clicks the “download” button. The data layer is downloaded and automatically added to the current HydroDesktop project.

Extensions: If the user is not logged in, HydroDesktop prompts the user to log in to HydroShare before searching for data layers. If the user does not have a HydroShare account, a link is provided to create one.

Approved for Implementation: [Yes or no. Who approved. When approved.]

Change Log: Cuyler Frisby 1/21/2014

<table>
<thead>
<tr>
<th>ID:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title:</td>
</tr>
<tr>
<td>Description:</td>
</tr>
</tbody>
</table>
HydroDesktop and logs in to HydroShare. The user uploads their project to HydroShare.

<table>
<thead>
<tr>
<th>Stakeholders:</th>
<th>The user that wishes to upload a HydroDesktop project to HydroShare. Any HydroShare users or organizations which may download and use the HydroDesktop project after it is uploaded to the HydroShare website.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actors:</td>
<td>The user that wishes to upload a HydroDesktop project to HydroShare. The HydroShare system manages access to HydroShare resources.</td>
</tr>
<tr>
<td>Trigger:</td>
<td>A user clicks the “Upload” button on the HydroShare ribbon in HydroDesktop.</td>
</tr>
<tr>
<td>Preconditions:</td>
<td>The user is logged into HydroShare. The user is authorized by the HydroShare system to upload resources.</td>
</tr>
<tr>
<td>Postconditions:</td>
<td>The HydroDesktop project files are uploaded to HydroShare and can be discovered and downloaded by other HydroShare users.</td>
</tr>
<tr>
<td>Main Success Scenario:</td>
<td>A user logs into the HydroShare website via the HydroShare ribbon in HydroDesktop. The user clicks the “upload” button located on the HydroShare ribbon in HydroDesktop. From a list of options, the user specifies that the current HydroDesktop project will be uploaded. All relevant files associated with the HydroDesktop project are zipped and uploaded to the HydroShare website.</td>
</tr>
<tr>
<td>Extensions:</td>
<td>If the user is not logged in, HydroDesktop prompts the user to log in to HydroShare before uploading any HydroShare resources. If the user does not have a HydroShare account, a link is provided to create one.</td>
</tr>
<tr>
<td>Approved for Implementation:</td>
<td>[Yes or no. Who approved. When approved.]</td>
</tr>
<tr>
<td>Change Log:</td>
<td>Cuyler Frisby 1/21/2014</td>
</tr>
</tbody>
</table>

### 2.2 Selecting a Programming Language

A first step in my research was to determine the programming language which would be used to develop the plugin. C# and Python were both considered. C# was selected because it is the language HydroDesktop is written in. “HydroDesktop includes a number of plug-ins developed by the HIS team, and also supports third party plug-ins that follow a standard, well-defined plug-in interface described at the project website” (Ames, Horsburgh, et al. 2012). Thus,
it would be easy to incorporate a new plugin written in C# into HydroDesktop. Python was selected for testing because it is a scripting language with very simple syntax. It was believed that providing a way to incorporate Python-based plugins into HydroDesktop would allow individuals with limited programming backgrounds to contribute to the development of HydroDesktop plugins. A prototype of the plugin was built using each language and the overall ease of programming as well as the capabilities of each language was assessed. The language which will be used to create the plugin was then chosen.

The pros and cons of developing in each language were assessed. With regards to programming syntax, Python was determined to be the simpler language to use. Python was considered because it would allow individuals with limited programming experience to contribute to the development of HydroDesktop plugins. However, Python-based plugins introduce further complications to the application architecture because HydroDesktop is written in C#. Thus, IronPython (“IronPython” n.d.), which is a version of Python compatible with the .NET infrastructure, was considered to bridge the gap between Python and C#, as shown in Figure 8 below.
Figure 8. Facilitating the Interaction Between HydroDesktop and Python via IronPython

The original version of the HydroShare ribbon in HydroDesktop consisted of a single button which launched a simple Python script, as shown below in Figure 9.

![Image of HydroDesktop ribbon with a button labeled Python]

Figure 9. Original Prototype HydroShare Ribbon in HydroDesktop

Launching a Python script from HydroDesktop using IronPython consisted of creating a Python “runtime” using IronPython, and then launching an external script in this runtime. The code to accomplish this is displayed below in a screenshot from Visual Studio in Figure 10.
Significant difficulty was encountered in getting this very simple tool to work. Ultimately, the difficulties caused by trying to integrate two different programming languages were considered to outweigh the ease of use of the Python programming language itself.

Although C# is considered more difficult to learn than Python, the simplicity of writing a C#-based plugin which interacts with the C#-based HydroDesktop application was determined to provide a more simple overall development experience than Python did.

2.3 Development of a Prototype HydroShare Ribbon for HydroDesktop

When the development of the HydroShare plugin for HydroDesktop began, the HydroShare website was in a beta-testing phase and was based on the Drupal (“Drupal” n.d.) web development framework. Working versions of the upload and download buttons were built and
tested using the web services which had been implemented. The prototype ribbon which was developed, featuring tools written in both C# and Python is shown below in Figure 11.

![Prototype HydroShare Ribbon in HydroDesktop](image1)

**Figure 11. Prototype HydroShare Ribbon in HydroDesktop**

The Python version of the HydroShare download tool was built using an open-source Python user interface library called wxPython (“wxPython” n.d.). A graphical interface creation application, wxFormBuilder (“wxFormBuilder” n.d.) was used for the creation of wxPython user forms. The Python-based prototype user form for the HydroShare download tool is shown below in Figure 12.

![Prototype Python-Based HydroShare Download Tool in HydroDesktop](image2)

**Figure 12. Prototype Python-Based HydroShare Download Tool in HydroDesktop**
These forms were then exported to Python scripts which were incorporated into the source code of HydroDesktop. The various event callers of the controls on the user forms were then connected to functions which were included on an additional Python script. For example, selecting a resource type from the combo box on the form will return all HydroShare resources which contain

The tool works by using URLs for HydroShare web services which perform functions such as list the available resources based on given search criteria or download a specified resource. Figure 13 below shows a code snippet of the URLs which were used for the prototype Python download tool. Once the new HydroShare web services are implemented, replacing these URLs with their new counterparts should restore functionality to this tool.

```python
class HydroShareDownloader():
    #base_url points to the export.php script that packages data files for download.

    #The url from which a list of available resources can be found
    list_url = "http://dev.hydroshare.org/?query_services/node.json&api-key=581d46dd"
```

**Figure 13. HydroShare Web Service URLs in Python Downloader Source Code**

However, in January of 2014, the Drupal development of HydroShare was scrapped in favor of a Django-based (“Django” n.d.) development. No web services are currently implemented in the new version of HydroShare, so the development of the HydroShare plugin has been delayed. Once the web services are implemented in the new HydroShare website, the development and testing of these tools will be continued.

### 2.4 Contributions to the HydroShare Website Development

In the meantime, my research focus has shifted toward contributing to the web development of the HydroShare website. This has included tasks such as defining what a
HydroDesktop project resource type will consist of, writing use cases for the HydroShare plugin (included in subsection 2.1), and the unit testing of HydroShare web services.
3 COMPLETED RESEARCH PAPER – CLIENT-SERVER HYDROLOGIC DATA EXCHANGE – MAKING THE (USE) CASE FOR DESKTOP APPLICATIONS

The following paper was completed to accompany a poster presentation to be given for the International Environmental Modelling and Software Society (iEMSs) 7th International Congress on Environmental Modelling and Software, to be held in San Diego, California on June 15 – 19, 2014. The contents of this paper were also presented as a 20-minute oral presentation at the American Water Resources Association (AWRA) 2014 Spring Specialty Conference on GIS and Water Resources at Snowbird, Utah on May 12, 2014. Following the paper, I have also included comments from the iEMSs peer review process. Suggestions for improving the paper were included in the version presented here.

3.1 BACKGROUND

The hydrologic sciences are hindered by a lack of efficient tools for the sharing and discovery of hydrologic resources and data (CUAHSI, 2010). The sharing of observational data, models, scripts, and analytical results will facilitate greater collaboration and easy access of data. Desktop applications, models, and tools are widely used in the hydrologic sciences because they enable editing and storing large amounts of data locally, using models that are intended for use on local computing resources. In many cases, however, it may be highly advantageous to use web-based resources that enable sharing of data, analyses, and results with a group of colleagues.
or the scientific community at large. This recognition has helped drive a movement to web-based applications in the hydrologic sciences; however, unique problems are presented by strictly web-based applications. For example, a deactivated or inaccessible server or ISP may inhibit the ability to access web applications. The approach presented here uses a mixed platform paradigm with a desktop application, HydroDesktop, linked to the HydroShare data and model sharing system via web services. This approach is expected to capitalize on the benefits of both off-line and on-line capabilities of the platforms.

Several projects and organizations are focused on the development of efficient data-sharing tools. DataONE is a multi-institutional, multinational, and interdisciplinary collaboration working to develop an organizational structure that will support the full information lifecycle of biological, ecological, and environmental data and tools to be used by researchers, educators, and the public at large (Allard, 2012). A focus of the DataONE infrastructure is to address the problems a researcher may find when he or she needs content from more than one data repository, each of which may be tailored to the needs of a particular domain or community of researchers (Allard, 2012). The Global Earth Observation System of Systems (GEOSS) aims to be a global and flexible network of content providers allowing decision makers to access a wide range of information at their desk. This ‘system of systems’ will proactively link together existing and planned observing systems around the world and support the development of new systems where gaps currently exist. It will promote common technical standards so that data from the thousands of different instruments can be combined into coherent data sets (“What is GEOSS?,” 2014).

One major NSF-funded initiative is the Consortium of Universities for the Advancement of Hydrologic Sciences, Inc. (CUAHSI). The main goal of CUAHSI is the advancement of water
science in the United States. This includes improving access to hydrologic data, information and models and facilitating interactions among the diverse water research community (CUAHSI, 2010). To this end, the CUAHSI Hydrologic Information System (HIS) is an internet-based system for sharing hydrologic data. It is comprised of databases and servers, connected through web services, to client applications, allowing for the publication, discovery and access of data (Horsburgh, Tarboton, et al. 2009). The primary client application of the HIS infrastructure is HydroDesktop, which gives users a convenient interface to access data, retrieves metadata from HIS Central and downloads data from HydroServers (Ames, Horsburgh, et al. 2012). CUAHSI’s new website, HydroShare.org seeks to provide an easy-to-use web-based interface for the sharing and discovery of hydrologic data. HydroShare users will be able to upload resources to the website which will then be discoverable by other users. This new paradigm for the sharing of hydrologic data will allow individuals with very different research objectives from around the world to collaborate and support one another (Tarboton, Idaszak, 2012).

Nearly infinite amounts of data are available in the many online catalogs and data repositories which already exist. Desktop GIS tools are capable of using and manipulating these data to produce useful conclusions and results. However, there is a dearth of desktop GIS tools which are connected to these online catalogs and data repositories. HydroDesktop fills this crucial niche for hydrologic and environmental applications. Expanding the capabilities of HydroDesktop to include support for HydroShare interaction will make it the flagship desktop client for downloading, manipulating, and sharing data from HydroShare. HydroDesktop users will be able to seamlessly share their data, models, and projects with one another.

HydroDesktop currently lacks a means whereby to interact with the new HydroShare website. The goal of this project is the development of a HydroShare plugin which will allow
the discovery and sharing of resources from within the HydroDesktop environment. Users will be able to search for shapefiles, data series, and projects and download these directly into their current HydroDesktop project. They will also be able to upload their current project or an individual data layer to the HydroShare website. The development of this plugin will greatly simplify the sharing and discovery of data, increasing the pace at which hydrologic science can progress. This paper will introduce the prototype user interaction tools for working with HydroShare once it is fully operational.

3.2 METHODS

The new HydroShare plugin for HydroDesktop will be written as a C# plugin for the CUAHSI HydroDesktop application. The HydroShare website, basin in Python, is currently being developed by a multi-institutional team of researchers associated with CUAHSI. A HydroDesktop project consists of multiple files which will be listed in detail below. These files will be uploadable to HydroDesktop via the HydroShare plugin. This “HydroDesktop Project” resource will inherit from the “composite resource” HydroShare resource type. This resource type allows separate resources to be linked to one another and downloaded as a single package. This aggregation of resources will be designed in accordance with the Open Archives Initiative Object Reuse and Exchange (OAI-ORE) specification (Lagoze, Van de Sompel, et al. 2008).

3.2.1 Defining a “HydroDesktop Project” resource type

Many diverse types of content will be shared on HydroShare, including data, models, model components, model packages, workflows, workflow packages, documents, and other types.
Because of this, a flexible data model is needed that enables HydroShare to structure content within the system, but that does not limit the potential contents of resources to be shared within HydroShare (Horsburgh, 2014). BagIt, a hierarchical file packaging format fits these requirements and is used to package all HydroShare content.

A “resource” is the discrete unit of digital content within HydroShare. Resources may be made up of a single content file or may be an aggregation of multiple content files, as in the case of a HydroDesktop project resource (Horsburgh, 2014).

To demonstrate which files make up a HydroDesktop project, a sample project using the “North America Map” template was created in HydroDesktop and saved in a new folder. The contents of this folder (i.e. the files which make up the HydroDesktop project) are shown below in Figure 14. The “search results” folder contains shapefiles for the site locations returned by a HydroDesktop search.

![Figure 14. Files Associated with a Sample HydroDesktop Project](image)
As shown in Figure 14, a typical HydroDesktop project will contain the following types of content:

- 1 HydroDesktop Project File (.dspx)
- 2 SQLite database files
- Several “base” shapefiles, each consisting of 4 individual files (.dbf, .prj, .shp, .shx)
- Several “Search Results” shapefiles

All of these files are unique to an individual HydroDesktop project except for the “base” shapefiles. These are used in the HydroDesktop template projects and therefore are present in many HydroDesktop projects which may potentially be uploaded to HydroShare. As shown in Figure 14, many of these shapefiles take up more disk space than any other files associated with the project. It is inefficient, both from a storage and transfer perspective, to store a copy of these files with each HydroDesktop project uploaded to HydroShare. Therefore, it is proposed that a single copy of commonly-used shapefiles be stored on the HydroShare website. Each project which uses these shapefiles that is uploaded to HydroShare will simply contain a link which references the copy of the file stored on the HydroShare servers. When the project is downloaded, the server will package all files which are contained in the project into a single BagIt file for download.

Storing commonly-used shapefiles on HydroShare brings the added benefit of users being able to search for and download these individual shapefiles for inclusion in their HydroDesktop or other GIS projects.

The BagIt file packaging format, which is used for all HydroShare resources, must contain the following required elements (Boyko, Kunze, et. al. 2012):

- Bag Declaration: bagit.txt
• Payload Directory: data/
• Payload Manifest: manifest-<alg>.txt

Additionally, a BagIt file may contain the following optional items:
• Tag Manifest: tagmanifest-<alg>.txt
• Bag Metadata: bag-info.txt
• Fetch File: fetch.txt
• Other Tag Files

The purpose of this paper is not to describe each of these items in detail, but to describe how the HydroDesktop Project content type will fit into the BagIt framework.

The payload directory contains the unique data (excluding “base” shapefiles) for the HydroDesktop project. Specifically, the HydroDesktop project file, SQLite database files, and “Search Results” shapefiles will be stored in the payload directory. The payload manifest is a tag file that lists these payload files.

The “Fetch File” is described in “The BagIt File Packaging Format” specification as follows:

“For reasons of efficiency, a bag MAY be sent with a list of files to be fetched and added to the payload… (Boyko, Kunze, et al. 2012)”

Thus, it is intended to have such a “Fetch.txt” file included in BagIt directories for HydroDesktop project resources. This text file will contain the URLs and Filenames for the “base” shapefiles on the HydroShare website. When a HydroDesktop project is downloaded, the server will search for these files and package them with the BagIt file to be downloaded. Figure 15 shows how the HydroDesktop project resource type will be stored in a BagIt file.
3.3 RESULTS

3.3.1 Server-Side Design (HydroServer Website)

HydroDesktop project resources will be available for download directly from the HydroShare website, without the use of HydroDesktop. The user will access these resources by navigating to the HydroShare website and logging into HydroShare. The user will then be able to search for and download the desired HydroDesktop project from HydroShare using the following steps:

1. A user logs into the HydroShare website
2. The user searches for the desired HydroDesktop project using the “Find Resources” section of the website
3. A list of available resources which meet the search criteria is returned
4. The user selects a HydroDesktop project from the list and is presented with information regarding the project.
5. The user downloads the HydroDesktop project by clicking the “download” button.
6. The HydroDesktop project files are downloaded and saved to a specified location on the user’s computer

If the user is not logged in, the website prompts the user to log in to HydroShare before searching for HydroDesktop projects. If the user does not have a HydroShare account, a link is provided to create one. A preview of the HydroDesktop project will be shown on the HydroShare website, including a screenshot of the project.

3.3.2 Client-side design (HydroDesktop)

HydroDesktop users will be able to download and upload complete HydroDesktop projects from HydroShare. A new “HydroShare” ribbon will be added to the HydroDesktop interface. A mockup of what this ribbon may look like is shown below in Figure 16.

![Figure 16. HydroShare Ribbon in HydroDesktop](image-url)
From within HydroDesktop, the user will search for and download the desired project from HydroShare and the project will appear in the HydroDesktop project window. The specific steps a user would take to download a complete HydroDesktop project are detailed below:

1. A user logs into HydroShare through the HydroShare ribbon in HydroDesktop
2. The user clicks on the “Download from HydroShare” button on the HydroShare ribbon in HydroDesktop and is presented with a search box and several filter options such as “resource type” and “keyword”
3. The user executes a search query and a list of available HydroShare resources which meet the specified criteria is returned
4. The user selects the desired HydroDesktop project and clicks the “download” button
5. The HydroDesktop project is downloaded and appears in the main window of HydroDesktop

The current prototype for the HydroShare download form is shown below in Figure 17.

![HydroShare Download Form in HydroDesktop](image.png)

Figure 17. HydroShare Download Form in HydroDesktop
The steps taken to upload a complete HydroDesktop project to HydroShare are given below:

1. A user logs into the HydroShare website via the HydroShare ribbon in HydroDesktop. The login screen which appears is shown below in Figure 18.

![Figure 18. HydroShare Login Screen in HydroDesktop](image)

2. The user clicks the “upload” button located on the HydroShare ribbon in HydroDesktop.

3. From a list of options, the user specifies that the current HydroDesktop project will be uploaded. The current prototype for the “resource type selector” form is shown below in Figure 19. Note that the “HydroDesktop Project” resource type is not yet included on this list.

![Figure 19. Resource Type Selector Form in HydroDesktop](image)

4. The necessary metadata fields are populated from the user’s HydroShare account, or are manually entered through a series of forms. The prototype metadata entry form is shown below in Figure 20.
Figure 20. Metadata Entry Form in HydroDesktop

5. All relevant files associated with the HydroDesktop project are zipped and uploaded to the HydroShare website.

If the user has not logged in to HydroShare, HydroDesktop prompts the user to log in to HydroShare before searching for or uploading HydroDesktop projects. If the user does not have a HydroShare account, a link will be provided to create one. If a HydroDesktop project was already open, HydroDesktop prompts the user to save it before a new project is downloaded.
3.4 CONCLUSION

The new HydroShare website, currently under development, will provide an easy way to share and discover hydrologic and climatic data, models, and projects. The goal of this project is the development of a HydroShare plugin which will allow the discovery and sharing of resources from within the HydroDesktop environment. The development of this plugin will greatly simplify the sharing and discovery of data, increasing the pace at which hydrologic science can progress.

Due to the current developmental status of the HydroShare website, the HydroDesktop plugin described in this paper has not yet been implemented. The user interfaces and client-side coding is nearly finished. However, until HydroShare reaches a more stable point in its development formal testing of these tools cannot be completed. Through active participation in the development of the HydroShare website, the HydroDesktop team at Brigham Young University will be able to work on both the client-side and server-side coding, increasing the level of cohesion and performance of this tool.

Once the HydroShare begins to take a more stable form, the HydroDesktop plugin for HydroShare will be tested and refined until it is ready for release.

3.5 iEMSs Conference Review Process

I wrote and submitted the above paper for the International Congress on Environmental Modelling and Software (iEMSs) 7th International Congress on Environmental Modelling and Software, to be held June 15 – 19, 2014 in San Diego, California. My paper was accepted for publication in the conference proceedings, after being evaluated by two reviewers for the conference. Their comments about the paper are included below.
PAPER: 316

TITLE: Client-Server Hydrologic Data Exchange – Making the (Use) Case for Desktop Applications

AUTHORS: Cuyler Frisby, Daniel Ames and Nicholas Martin

-------- REVIEW 1 --------

This is a very well-written paper that addresses a highly-relevant development in a major software system, impacting a wide array of environmental scientists and enhancing their ability to share data. The implementation of the new data sharing feature within CUAHSI requires that issues of data deduplication, file size, and general data organization be addressed - all of which are discussed quite nicely. These are all non-trivial points that many developers of environmental data systems fail to address or plan to incorporate, leading to dead-end systems (i.e. "abandonware"). Bringing these issues - as well as mechanisms for resolution - to the attention of the general research community is of great benefit.

The paper itself is extremely well-organized and well-written, making its points clearly and concisely. The illustrations provided are also appropriate and clearly readable, adding nicely to the text. Formatting guidelines are heeded and appropriate citations are present.
This is an excellent paper that fits the requirements of the conference perfectly. The authors and their associates are clearly innovative and dedicated to contributing to the development of the larger global research community.

-------- REVIEW 2 --------

This paper is well written and appears to be complete. It also seems to conform to all of the formatting requirements (as would be expected since one of the authors is also an editor :)). I find little to criticize in this paper and didn't notice anything that was missing in terms of explanation or background or references. However, the Conclusions section seems a little bit "over-apologetic” about the status of the HydroShare project. Although beyond the scope of the paper, I would be interested to see ways that additional metadata about models and data sets at the level of exchange variables and underlying assumptions could also be collected from users, as this is an active area of research.

These suggestions were implemented into the final form of that paper, and its contents will be presented at the conference as a poster presentation. I also had the opportunity to present the contents of this paper at the U.S. Army Corps of Engineers Engineer Research and Development Center Coastal Hydraulics Laboratory in Vicksburg, Mississippi as part of a job interview. My formal presentation to fulfill the requirements for my Master’s project took place on May 12, 2014 at the American Water Resources Association (AWRA) Spring Specialty Conference on Water Resources and GIS in Snowbird, Utah.
4 CONCLUSIONS

The new HydroShare website, currently under development, will provide an easy way to share and discover hydrologic and climatic data, models, and projects. The goal of this project is the development of a HydroShare plugin which will allow the discovery and sharing of resources from within the HydroDesktop environment. The development of this plugin will greatly simplify the sharing and discovery of data, increasing the pace at which hydrologic science can progress.

Due to the current developmental status of the HydroShare website, the HydroDesktop plugin described in this paper has not yet been implemented. Prototypes of the user interfaces and client-side coding is nearly finished. Through active participation in the development of the HydroShare website, the HydroDesktop team at Brigham Young University is currently working on both the client-side and server-side coding, increasing the level of cohesion and performance of this tool. Once the necessary web services are implemented on HydroShare, the development of the HydroDesktop plugin can continue.

As the amount of material on HydroShare increases, it will become necessary to implement additional filtering tools for the download tool on the HydroShare ribbon in HydroDesktop. The current prototype version only allows filtering by resource type. Additional filters that may be of benefit include keyword, title, creator, date published, and so on.
After the plugin is built, user testing and performance testing will be completed to measure and improve the usability and performance of the software.
REFERENCES


