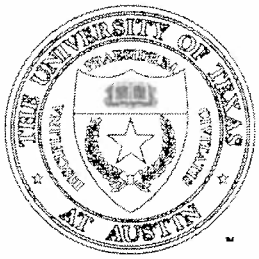




# THE TEXAS HYDROLOGIC INFORMATION SYSTEM

## FINAL REPORT

**8/31/2009**

**Contract: 0904830897**

**Prepared by:**
**Tim Whiteaker (twhit@mail.utexas.edu)**
**Center for Research in Water Resources**
**The University of Texas at Austin**



# THE TEXAS HYDROLOGIC INFORMATION SYSTEM

## FINAL REPORT

**8/31/2009**

**Contract: 0904830897**

**Prepared by:**
**Tim Whiteaker (twhit@mail.utexas.edu)**
**Center for Research in Water Resources**
**The University of Texas at Austin**

## Executive Summary

A project was undertaken by the current Principal Investigator, Dr. David Maidment of The University of Texas at Austin Center for Research in Water Resources (CRWR), with assistance from Dr. Tim Whiteaker and Bryan Enslein, in FY2008 to develop a statewide hydrologic information system (HIS) for Texas and to populate the prototype system with example available datasets for select priority basins. As a result, the prototype Texas HIS saw improvements to its functionality, usage, and depth and breadth of data incorporated. Significant accomplishments include conducting two HIS training workshops in Austin, working with TNRIS to develop the TxHIS Viewer, populating and publishing databases of water observations data, and developing strategies for registering published datasets online.

# Table of Contents



# List of Tables

# INTRODUCTION

A project was undertaken by the Center for Research in Water Resources (CRWR) at The University of Texas at Austin to develop a statewide hydrologic information system (HIS) for Texas and to populate the prototype system with example available datasets for select priority basins. Bryan Enslein, a (now graduated) Masters engineering student at The University of Texas, was the primary student working on this project, with technical supervision from Dr. Tim Whiteaker and overall coordination by Dr. David Maidment.

As a result of this project, the Texas HIS saw improvements to its functionality, usage, and depth and breadth of data incorporated. Significant accomplishments include conducting two HIS training workshops in Austin, working with TNRIS to develop the TxHIS Viewer, populating and publishing databases of water observations data, and developing strategies for registering published datasets online.

This report documents work flow processes, methods researched and technologies used in the project. The report then summarizes accomplishments of the project.

For quick reference, the following websites have information pertinent to this project:

Details of methodologies used for building Texas HIS:
http://www.twdb.state.tx.us/RWPG/rpgm_rpts/0704830777_TexasHIS.pdf

Inventory of datasets published online with Texas HIS:
http://data.crwr.utexas.edu/

Texas HIS Viewer:
http://waterdatafortexas.org/

Prototype HydroPortal for data publication:
http://crwr-idis.crwr.utexas.edu:8080/GPT9/

HIS homepage:
http://his.cuahsi.org/

The scope of work is shown below.

---

**SCOPE OF WORK**

David R. Maidment, Principal Investigator

Access to relevant data is of paramount importance in assessing, conserving, and developing the water resources of the State of Texas. A project was undertaken by the current Principal Investigator, Dr. David Maidment of The University of Texas at Austin Center for Research in Water Resources (CRWR) in FY2008 to develop a statewide hydrologic information system (HIS) for Texas and to populate the prototype system with example available datasets for select priority basins. The proposed project builds

upon the prototype Texas HIS to increase its functionality, usage, and depth and breadth of data incorporated via the following five tasks.

CRWR personnel will:

1) Populate Texas HIS with data sources listed in Appendix A, including water levels and water quality observations from Texas Water Development Board groundwater. Create either Observations Data Model web services or Hybrid Metadata Wrapper web services, whichever is best suited, for each data source. Once Texas HIS has been populated with data from all the data sources mentioned above, continue to broaden the range of data sharing partners by considering them as potential dataset contributors on a case by case basis, which will be subject to Texas Water Development Board approval for inclusion into Texas HIS.

2) Provide training programs and documentation for Texas HIS. Two workshops will be organized and offered, covering an introduction to Texas HIS, data access, and detailed instruction regarding how to provide and load data. The target audience includes data management personnel from TWDB, TNRIS, TCEQ, and other relevant federal, state, and local organizations.

3) Coordinate development efforts with TNRIS and TWDB to ensure compatibility with parallel TNRIS efforts for data access, data viewing, and geographic data integration. Assist TWDB and TNRIS personnel in their data loading, aggregating, and querying efforts. This would include assistance with the installation, deployment and testing of a copy of UT-CRWR's central registration and publication system at TNRIS. Hold conference calls periodically to discuss progress and resolve issues.

4) Create automated procedures which consist of computer programs and or scripts to periodically append data and metadata from data sources to ODM databases and metadata catalogs, respectively within the Texas HIS.

5) Provide quarterly progress reports. Provide a final report documenting work flow processes which include instructional sections for proper use of programs and scripts. Also, provide documentation of methods researched and technologies used in the final report.

Proposed Project Period: 9/1/08 to 8/31/09

Proposed cost: $95,480

CENTER FOR RESEARCH IN WATER RESOURCES
Texas HIS

| | FY 08-09<br>9/1/08 to 8/31/09 |
|---|---|
| **Salaries and Wages:** | |
| **David Maidment - Prinicipal Investigator**<br>0.5 month summer salary | 7,767 |
| **Tim Whiteaker**<br>1.25 calendar month | 7,375 |
| **1 Graduate Research Assistant**<br>12 months at 50% time | 19,200 |
| **Network Analyst/ITG**<br>2.5 calendar months salary | 11,415 |
| **Administrative Support**<br>15% Total salaries | 6,863 |
| **Fringe @ 28%:** | 14,734 |
| **Maintenance and Operations:**<br>(Possible expenditures include consumable materials & supplies, services, telecommunications duplicating, software, non-capitalized equipment, research materials, laboratory supplies, computers, computer components and equipment repair) | 3,173 |
| **Travel:**<br>(1 trip ESRI User's Conference for graduate research assistant and national AGU meeting in San Francisco) | 2,500 |
| **Tuition:**<br>Annual tuition cost based on historical data | 10,000 |
| **Total Direct** | **83,026** |
| **Indirect costs @ 15% TDC** | **12,454** |
| **Total Budget** | **95,480** |

**Tasks and Expense Category Budgets**
**Texas Hydrologic Information System**

| Task | Description | Amount |
|---|---|---|
| 1. | Populate Texas HIS with water data services and create web services | $25,599 |
| 2 | Provide training programs and workshops for target audiences | $19,238 |
| 3. | Coordinate development efforts with TNRIS and TWDB for compatibility for data access. | $22,587 |
| 4. | Create automated procedures which consists of | $22445 |
| 5. | Quarterly Progress Reports , Final Instructional documentation of methods and technologies | $19,095 |
| **Total** | | **$95,480** |
| | | |

# IMPLEMENTING TEXAS HIS

This section documents work flow processes, methods researched and technologies used in the project. There are two main tasks pertinent to this discussion: Data publication with WaterOneFlow web services, and data registration. These tasks are summarized in the text below. The Master's thesis of Bryan Enslein, the main student working on the Texas HIS project, describes these tasks in tremendous detail. The reader is referred to the thesis if additional information beyond that presented in this report is desired. The thesis was delivered to TWDB as partial fulfillment of contract 070483077, and is also available online at http://www.twdb.state.tx.us/RWPG/rpgm_rpts/0704830777_TexasHIS.pdf.

## DATA PUBLICATION WITH WATERONEFLOW WEB SERVICES

From agencies to academics, Texas has numerous sources of quality datasets ripe for publication. However, each dataset tends to have its own format and its own access mechanism (if any). The CUAHSI WaterOneFlow web service is a standard design for machine-to-machine access of hydrologic data. Publishing data with Texas HIS means using WaterOneFlow web services as the standard framework for communication of water data. To describe methodologies used towards this end, this document "works backwards" by starting with the WaterOneFlow web service and using experiences learned to guide development of a database of water observations. Three strategies are used for publishing data with WaterOneFlow:

(A) Prepare a database that supports the Observations Data Model (ODM), load data into this database, and wrap the database with the ODM WaterOneFlow service available as a free download from http://his.cuahsi.org/wofws.html.
(B) Load only site and variable data into an Observations Data Model (ODM) database and wrap the database with the ODM WaterOneFlow service. Set this service up as private. Wrap the private service with a public and custom-written WaterOneFlow web service that uses the private service to retrieve site and variable information from the underlying ODM database, and additional code to parse time series data from a separate location. This is called the Hybrid WaterOneFlow Web Service approach.
(C) Write a fully customized WaterOneFlow web service independent of ODM.

Option (A) is chosen as the best approach for Texas HIS for the following reasons:

- Generally, the time taken to implement (A) is much less than (B) or (C).
- No custom programming of web services is required.
- Not only are data published in a consistent way, but also the underlying databases use the same schema, further enabling sharing of data among Texas institutions.
- The ODM is supported by a suite of free tools for data loading, query, editing, and visualization, available from the HIS website at http://his.cuahsi.org/.

Once the decision has been made to use ODM, the next step is to determine how best to prepare the ODM database. For streaming data arriving from sensors operating in the field, the best approach is to use the Streaming Data Loader, documented at http://his.cuahsi.org/odmsdl.html. For existing SQL Server databases, write table views to create an ODM appearance on the data without changing the underlying data structure. This eliminates the need to reprogram any existing tools operating on the database, yet also supports the ODM

WaterOneFlow web service. Documentation on table views is available with SQL Server (in particular, view help on creating table views with SQL Server Management Studio.)

For archives of data files such as text files, two methods of loading data were investigated: SQL Server Integration Services (SSIS) and the ODM Data Loader (documented at http://his.cuahsi.org/odmdataloader.html). While SSIS provides more power and flexibility than the ODM Data Loader, it also requires far more training in both SSIS and ODM. From experience at CRWR, the ODM Data Loader can be learned in 10% of the time as SSIS, and also does not require advanced knowledge of ODM because the ODM Data Loader has quality control checks for ODM built into it. Therefore, the ODM Data Loader is the preferred software when loading data from archives of files.

Armed with these insights, the methodology for publishing hydrologic observations data with WaterOneFlow becomes clear:

1. Prepare an ODM database that contains the data to be published.
    a. For streaming data, use the Streaming Data Loader to load data.
    b. For files of data (such as archived data), use the ODM Data Loader to load data.
    c. For existing SQL Server databases, write ODM table views to give the data the appearance of ODM.
2. Install the ODM WaterOneFlow web service from the HIS website.
3. Configure the service to use the ODM database as its source of data.

By following this methodology, a given dataset becomes accessible in a standardized way just like other datasets published in Texas HIS. Instructions for using each piece of HIS software are available on the HIS website (http://his.cuahsi.org/). Details of the overall methodology can be found in Bryan Enslein's Master's thesis (http://www.twdb.state.tx.us/RWPG/rpgm_rpts/0704830777_TexasHIS.pdf).

## Registering a Texas HIS Data Service

Once a WaterOneFlow web service for Texas data has been created, it needs to be registered with Texas HIS so that others can discover and use it. It is envisioned that TWDB will host a final registration system as part of http://waterdatafortexas.org/. For now, a website cataloging published data services is hosted at http://data.crwr.utexas.edu/. Because all data services have been developed by CRWR or HIS affiliates, the "registration" process has simply been for the CRWR team to:

1. Manually create a new web page on the web site for the data service.
2. Use ArcGIS Server to create a map service showing site locations. The service is published as an ArcGIS map service, and optionally as a Web Feature Service (WFS) and Web Mapping Service (WMS).

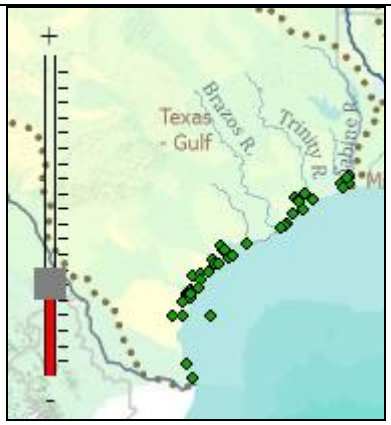
The procedure to register a service will change as http://waterdatafortexas.org/ evolves.

To browse for a given data service:

1. In a web browser, navigate to http://data.crwr.utexas.edu/.
2. Click the link for "WaterML Data Services."
3. In the list of services, click the service of interest.
4. On the service page, click to interact with the map of site locations or download supporting files.

The web page for a given service includes the information shown in Table 1:

**Table 1 Items included on web page for registered WaterOneFlow service**

| Item | Example |
|---|---|
| Data Source | Texas Water Development Board |
| Description | TWDB coastal water quality data |
| WSDL Location | http://his.crwr.utexas.edu/TWDB_Sondes/cuahsi_1_0.asmx?WSDL |
| No. of Sites | 56 |
| No. of Variables | 5 |
| Variables | * pH<br>* Temperature<br>* Conductivity<br>* Salinity<br>* Dissolved Oxygen |
| No. of Values | 117,840 |
| Map of site locations |  |
| Link to GIS Services | http://129.116.104.172/ArcGIS/rest/services/TxHIS/TWDB_CoastalWQ_Sites/MapServer |
| Related Files | links to zip files of items such as site locations as KMZ, site location as shapefile, HydroExcel populated with data from the given data service |

Towards the end of the contract period, a new technology emerged called the Geoportal Extension. This software enables a server to host a portal for registered items indexed by metadata and geographic extent. Users can search for items via an interactive map or by keyword search. The search process can be automated via the Geoportal Extension's REST interface, which is built to be conformal to the Open Geospatial Consortium's Catalog Server (for the Web). The extension can also be customized, and a customization for hydrology applications called HydroPortal is currently under development by ESRI. A prototype HydroPortal for publication of Texas water data services is at http://crwr-idis.crwr.utexas.edu:8080/GPT9/. Future work with Texas HIS will involved evaluation of HydroPortal as a viable registration and publication path for water data services, and potential implementation at TNRIS.

# Accomplishments and Deliverables

This section details how deliverables as defined in the scope of work were met.

## Populate Texas HIS with Data Sources from a Priority List

Data sources identified as the highest priority for inclusion in Texas HIS were identified in Appendix A of the Scope of Work. The data sources for which WaterOneFlow services have been developed are shown in Table 2. The data sources with Texas scope are indexed at http://data.crwr.utexas.edu/, while those with national scope (e.g., the USGS service) are indexed at the National HIS Central at http://hiscentral.cuahsi.org/.

**Table 2 WaterOneFlow services for prioritized data sources**

| Data Source | WaterOneFlow Service URL |
|---|---|
| TWDB Datasondes | http://his.crwr.utexas.edu/TWDB_Sondes/cuahsi_1_0.asmx |
| TWDB ADCP | http://his.crwr.utexas.edu/TWDB_ADCP/cuahsi_1_0.asmx |
| TWDB Water Quality | http://his.crwr.utexas.edu/TWDB_Quality/cuahsi_1_0.asmx |
| TWDB Tides | http://his.crwr.utexas.edu/TWDB_Tides/cuahsi_1_0.asmx |
| TWDB Wind | http://his.crwr.utexas.edu/TWDB_Wind/cuahsi_1_0.asmx |
| Texas Evaporation | http://his.crwr.utexas.edu/TxEvap/cuahsi_1_0.asmx |
| USGS Streamflow | http://river.sdsc.edu/wateroneflow/NWIS/DailyValues.asmx<br>http://river.sdsc.edu/wateroneflow/NWIS/UnitValues.asmx |
| TCOON | http://his.crwr.utexas.edu/tcoonts/tcoon.asmx |
| TPWD | http://his.crwr.utexas.edu/TPWD/cuahsi_1_0.asmx |
| TCEQ TRACS | http://his.crwr.utexas.edu/TRACS/cuahsi_1_0.asmx |
| NWS Rainfall | http://his.crwr.utexas.edu/nwsmpe/cuahsi_1_0.asmx |

While developing a system where all of these data sources facilitate online access to their data in a standard way can be deemed as an accomplishment, there is an additional step that can be performed to add value to this result. A data curator can identify time series of interest related to a given theme from all Texas data sources. For example, all of the time series related to salinity can be identified, and descriptive information about those time series can be stored in a single table called a **Theme**. This table has basic information about the time series such as the location of the site of measurement and the name of the variable being measured. But it also contains all of the information needed to make a WaterOneFlow web service request to download the time series. Thus, if one were interested in getting salinity data for Texas, one would not have to search each individual Texas HIS web service for salinity data. Rather, one could simply access the Salinity theme and then download data from the theme as desired.

The data curator can add even more value by compiling all of the time series data for the theme into a single database, and posting the database online for easy access. This saves the user the task of making each individual time series request for data. A single request for a zipped database is much faster than thousands of request for WaterML responses. Of course, this approach works best with static datasets or with the understanding that the theme is a snapshot of the data in time. A mechanism for updating themes real-time datasets has yet to be researched.

## Provide Training Programs and Documentation for Texas HIS

Texas HIS workshops were conducted in October, 2008, and May, 2009, in Austin, TX. The May workshop included hands-on training with printed course materials and sample data that walks the user through the data publication process. These materials were submitted to TWDB following the workshop. Additional electronic copies can be provided upon request.

## Coordinate Development Efforts with TNRIS and TWDB for Web Data Access

CRWR has worked with Richard Wade of TNRIS to refine the **Texas HIS Viewer**. The viewer can now display site locations for WaterOneFlow web services hosted by CRWR, identify time series for observation sites, and plot graphs of one or more time series variables.

## Create Programs or Scripts to Support Data Updates

The conclusion from investigations into this aspect of the scope of work is to rely on existing software freely available for working with HIS.

- To update the database with data streamed from sensors, use the **Streaming Data Loader**.
- To update the database with data from an ordinary text file, use the **ODM Data Loader**.
- If the database includes ODM table views, then update the database with whatever mechanism is already in place for database updates. Table views are dynamic and will be updated to reflect the most recent information in the database.

## Provide Quarterly Progress Reports and a Final Report

Quarterly reports were provided as outlined in the scope of work. This document serves as the final report.

# Conclusions and Recommendations

Through the work conducted for this project, water data for Texas are now published online in a standard way, including some datasets that were not even available online before this effort began. Aside from the obvious task of bringing more datasets into Texas HIS, additional recommendations for future work were elucidated by undertaking this project:

- Themes (curated sets of data related to a particular topic across one or more Texas data services) should be investigated further to determine the proper format (schema), to develop automatic theme harvesting software, and to test feasibility of supporting themes for real-time data,
- The HydroPortal should be evaluated as a mechanism for indexing registered Texas HIS web services and themes.
- Texas agencies and academic institutions should continue to be educated about Texas HIS to encourage participation.