OPEN WATER

# Foundational Data Services for Water Resources Intelligence within the NWS Office of Water Prediction

John A. McEnery[1], Nathan R. Swain[2], Alan J. Harmon[3]

[1]*UCAR CPAESS - NOAA/NWS/OWP National Water Center*
[2]*Aquaveo & Lynker Technologies - NOAA/NWS/OWP National Water Center*
[3]*NOAA/NWS/OWP National Water Center*

**ABSTRACT**

The National Weather Service (NWS) Office of Water Prediction (OWP) is building an extensible web services framework for discovery and access of select hydrologic and meteorologic data, metadata, and geospatial information. The services framework will be implemented using community data exchange and format standards to provide an interoperable platform that will support an ecosystem of application development both within OWP and by external partners in the public, private, and academic sectors. The initial implementation of the services framework will encompass three essential classes of dissemination for hydrologic data: gridded data services, time-series services, and geospatial services. Development priorities are based on the data services needed in evaluation, modeling, and dissemination supporting the interoperability goals of Integrated Water Resources Science and Services (IWRSS) consortium partners and the development of the National Water Model (NWM). The services framework will integrate currently fragmented water information into a connected hydroinformatic system for water data that leverages existing applications, infrastructure and capabilities for the support of national water resources objectives.

*Keywords*
hydroinformatics, data services, geospatial, open data, water resources

## 1.0 Introduction

The Water Resources Data Services (WRDS) project, within the National Water Center (NWC), is developing capabilities for interoperable access of data needed to support improved hydrologic forecasting. Data access is essential to many of the projects currently in progress at the NWC. Timely exchange of informational sets in a common recognizable format is crucial for operation of the National Water Model and its extended achievements, as well as to streamline dissemination of analytics produced with these new modeling tools. One of the foundations of the NWC is its role as a catalyst for the Integrated Water Resources Science and Services (IWRSS) partnership (NOAA, 2009). A key tenet of the IWRSS partnership is to integrate information delivery and simplify access to data. As the WRDS project establishes new systems for data access and dissemination, it will directly advance the goals of the IWRSS initiative for integrated science and services across organizations and stakeholders. These new services will also help in complying with Presidential Executive Order 13642 and the Open Water Data Initiative (Blodgett et al., 2016).

## 2.0 Project Development

The initial focus of the WRDS development project was assessment of functional expections and potential solution strategies for the system. WRDS undertook the effort to analyze system requirements, user needs and community exchange standards in order to plan, design, develop and implement an enterprise-wide hydroinformatics data service. This included the synthesis of information from a prior planning effort for the NWC referred to as the Water Forecast Improvement Preparatory Project (WFIPP) which had gathered a volume of participant input calling for state-of-the-art information and communications technologies to transform the accessibility and usefulness of hydrologic informational resources.

### 2.1 Requirements

Through assessment of the project scope it was determined that this data service should collect and retain data, metadata and quality control information for current and historic values including some or all of the following; streamflow observations; remotely sensed observations; water management observations; gridded meteorological and hydrologic analysis; model based analysis and forecasts; and geospatial information. Examination of existing WFIPP planning documents, along with service asssessment reports and IWRSS guidelines resulted in a core set of requirements for the WRDS project. Key requirements include the following:

- Create central hydrologic information system of the NWC for cataloging, discovery and dissemination of diverse informational data sets of record.
- Determine and implement appropriate data ontology, catalog system, data service protocols and integrate suitable APIs for a scalable, state-of-the-art enterprise system
- Create and maintain a catalog of data described by extent in time, space and associated identifiable attributes, and make it available to a requesting party for discovery of data contents with functionality for individual, recurrent or automated access.

### 2.2 Design Factors and Development Approach

In addition to the project requirements defined above, the evaluation of potential design solutions began with specification of several key criteria and adoption of a general approach. Central to this was project guidance adopted from a working group of the IWRSS initiative. The IWRSS Interoperability and Data Synchronization (IDS) team proposed policies, techniques and systems for increased efficiency and common understanding of critical data (Briar et al., 2015). IDS goals are accomplished through use of standardized methods for data exchange by which individual systems are better able to communicate with one another and through adopting community standards in data and metadata to enable coherency across system processes. This is a platform agnostic approach toward making water resources data available with semantic consistency using machine readable protocols, as stated in the IWRSS System Interoperability and Data Synchronization Report (2013). These IDS concepts have been carried forward into the WRDS project development.

A primary guideline of the WRDS framework for sharing water information is the use of a pull rather than push data exchange process. This paradigm is based upon the need to "Share information, not systems" (Briar et al., 2015). This does not permit external users to push data into holdings of an agency's system, directly access the internal elements of the system or modify the contents and software of that system. Propagation of information occurs through transactions initiated by an external request to which responsive data is returned. Interactions between systems require each of the systems to know only the external interface of the other; specifically, how to communicate data requests and acknowledgements with the other system and format of the data being transferred. With each of these exchanges, conformity of standards and the alignment of consistent information about data provenance ensures that authoritative data are propagated through the enterprise. This "pull" methodology provides more secure, reliable and fault-tolerant data synchronization between agency computer systems.

To avoid risk of inter-agency operations being adversely impacted by conflicting, non-authoritative data propagating through the work flow, as well as duplication of system storage, it is the intention of the WRDS project to avoid to the fullest extent possible duplication of data delivered through other NWS platforms or through other IWRSS partner agencies data services. WRDS will provide a means for hosting "record" sets of original official information.

The WRDS project is being built in an incremental fashion intended to allow the system to scale with development and to allow for adaption to arising needs. The framework is meant to be extensible for

current and future use with each phase of development implementing new functionality or enhancements to existing functionality. Each successive iteration implements capability that is directed progressively towards building a comprehensive suite of services. This offers a design approach that is more sustainable and more responsive to advances in information technology and water resources science. The development approach is to leverage existing software capabilities and data management strategies where appropriate to meet prioritized service requirements. Fortunately, there has been a large endeavor within the hydrologic community to develop tools and software that accomplish many of the requisite operations for providing machine-to-machine interoperability with semantic meaning in a distributed environment of heterogeneous datasets. That body of open source development can be drawn upon for adaption either directly or conceptually into the WRDS framework. The initial set of functional services will continue to be augmented with new capabilities which integrate water information tools into a connected water information framework.

To fully support IWRSS objectives, it is critical that the WRDS project adopt community and industry standards for metadata and ontological conventions, data formats and exchange protocols. The Open Geospatial Consortium (OGC), Unidata and the Consortium of Universities for the Advancement of Hydrologic Science, Incorporated (CUAHSI) are each communities of practice that support significant data models and tools. From these communities and others, the WRDS project has recognized and is proceeding with technical design conforming to several key standards which include the following;

- WaterML for the dissemination of time-series information (Zaslavsky et al., 2007)
- CF compliant netCDF format gridded data files (Rew et al., 2006)
- Representational State Transfer (REST) data service protocols
- OGC compliant web-services for geospatial data.

The WRDS development project has also embraced another emerging hydrologic framework which is the NHDPlus geospatial dataset (McKay et al., 2012). NHDPlus is used as the spatial scheme to which the National Water Model is referenced. To insure the ability of WRDS to broker NWM information, the NHDPlus common identifiers (ComIDs) are included within WRDS metadata and cross-referenced to other spatial references; including latitude/longitude, NWS Location Identifier (NWSLI), USGS NWIS ID and USACE Site ID.

### 3.0 Organization of Services and Implementation

To accomplish the objectives defined for WRDS, the system needs to deliver a broad suite of functionalities and meet disparate informational goals with dissimilar data sets. As the WRDS project has moved into the design phase, development is focused upon the initial components of a coordinated suite of services. WRDS is currently implementing the following three essential types of service; gridded, time-series and geospatial.

Gridded data services: Typically work with directories of saved files in which a field of values for one or more physical elements are represented at equidistant locations distributed uniformly over a spatial domain.

Time-series services: Access to the values of a physical element at a discreet location or set of locations for a defined span of time. A series of values is returned to a client in a structured or unstructured text format using a protocol such as REST.

Geospatial services: Those which typically use a map server to generate and deliver information to a client in the form of georeferenced map images identifying hydrologic values formatted in compliant standards through an interface such as Web Map Services.

Thus far, WRDS has undertaken four distinct implementation projects. Recently completed are projects that: (1) have established a data service for gridded Precipitation Frequency Data Server (PFDS) NOAA Atlas 14 information, and (2) demonstrated a WaterML formatted time-series data service within the NWC for access to select forecast variables from the NWS River Forecast Centers. In progress are projects that will: (3) provide access to NWS Ensemble Streamflow Prediction (ESP) water supply forecasts through an online toolset and (4) provide data service for access to an array of information gathered from collaborating authorities. Also, foundational geospatial capabilities are being established to provide map services for the dissemination of the NWM spatial data through OGC and Esri REST compatible standards.

### 3.1 THREDDS Gridded Data Service

The first development under WRDS was the implementation of a service providing access to gridded data from a national mosaic of Precipitation Frequency Data Server (PFDS) NOAA Atlas 14 information. This

service was developed in conjunction with the OWP Hydrometeorological Design Studies Center (HDSC) group. The purpose was to extend existing web service capabilities of the HDSC group by adding gridded data service functionality. This is accomplished while leveraging existing software and community standards through use of a Unidata THREDDS Data Server. The information is hydrometeorological data contained within the respective volumes of PFDS from the NOAA Atlas 14. The individual volumes were assembled into a single netCDF CF-4 compliant file. This provides access to user chosen spatial selections of gridded values from a nationally mosaicked representation of NOAA Atlas 14 precipitation frequency values continuous across state boundaries. This service is active and can be located within the homepage for PFDS.

### 3.2 RFC Time-series Data Service

The purpose of this implementation was to provide WaterML time-series service within the NWC for access to select hydrologic variables from the NWS River Forecast Centers (RFCs). Each of the 13 RFCs transmit information daily for selected Advanced Hydrologic Prediction Service (AHPS) (McEnery et al., 2005) forecast locations; approximately 15,500 points. The specific parameters that are maintained include River Discharge, Runoff Inflow to Channel, Precipitation Areal Mean and Air Temperature Areal Mean. This implementation leverages existing software and community standards to provide data returned in WaterML2.0 format using a FEWS Data Server with the REST protocol. This service is now active. These time-series are made available within the NWC firewall system for internal use by OWP teams.

### 3.3 Water Resources Monitor and Outlook

This implementation is underway to support the extended development of an effort that began among a number of western region offices, called the National Water Resources Monitoring and Outlook Webpage (WRMO). An online toolset is being developed by the Colorado Basin River Forecast Center to provide access to NWS Ensemble Streamflow Prediction water supply forecasts. Ensemble forecast of streamflow is accomplished by using a 30-year history of temperature and precipitation records to compute hydrologic responses that would result given today's current state of conditions. This produces a set of 30 "traces "or possible hydrologic outcomes which can be compared and analyzed. WRMO users are presented a web map to navigate for selection of a desired forecast location. Upon selection of a location, the ensemble members are then retrieved by a machine-to-machine JSON exchange from the WRMO server. The WRMO web page provides on the client side an interactive toolset for plotting and analysis. The WRDS team has coordinated with CBRFC for configuration of a prototype development platform that has been installed within NWC.

### 3.4 Multiple Outreach Acquisition Data Service (MOADS)

The purpose of this implementation project is to provide access to an array of information gathered from collaborating authorities which will include state, regional and external federal sources of observation data. A specific priority for the Multiple Outreach Acquisition Data Service (MOADS) is acquisition of observed reservoir discharges. The source from which USACE reservoir observations will be accessed is the USACE RESTful API for Data Retrieval (RADAR). The WRDS design approach for MOADS is to again leverage existing software and community standards to acquire and host internally a set of pertinent hydrologic time-series data from multiple external partners. The MOADS system consists of a Postgres database with a CUAHSI Observations Data Model (Horsburgh et al., 2016) schema for hydrologic information. Time-series data is served using a WOFpy (Pothina & Wilson, 2011) web service API for WaterML with both REST and SOAP endpoints. The Unidata Local Data Manager (LDM) connection installed on an LDAD server at NWC will be used as a method for transmission of data into the NWC. Data curated within the MOADS system are maintained with spatial references based upon NHDPlus ComIDs. Additionally, metadata for all MOADS observation points include the nearest NWSLI and their native agency site ID.

### 3.5 Foundational Geospatial Capabilities

The NWS Geospatial Intelligence Division is investigating the capabilities of ArcGIS Enterprise 10.5 as the means for implementing geospatial web services for WRDS. ArcGIS Enterprise is implemented using a distributed architecture that allows it to scale to meet services demand. Geospatial data can be hosted as ArcGIS

REST Map and Features services as well as open formats such as OGC-WMS and OGC-WFS. Services can be readily consumed by desktop GIS and web applications to support NWC partners. In addition to providing a platform for data services, ArcGIS Enterprise includes a Portal for ArcGIS, which can be used to build interactive web applications on top of data services hosted by the system. This provides a platform for quickly building dynamic visualizations of NWM output.

## 4.0 Conclusions

The Water Resources Data Service of the National Weather Service (NWS) Office of Water Prediction (OWP) is developing a hydroinformatic framework for discovery and access to water resources data. This effort is leveraging community data exchange protocols, standardized data formats and shared software solutions to produce an interoperable platform for use within OWP and by external partners in the government, public, private, and academic sectors. It is also important to recognize that the evolving nature of information technology and water science requires a design approach that is both iterative and adaptive. Therefore, adoption of an incremental development approach offers greater long term sustainability for the system and greater responsiveness to arising data needs. These efforts are critical to provide a coherent hydroinformatic system for water data that leverages existing applications, standards and capabilities for the support of national water resources objectives and the interoperability goals of IWRSS partners.

## References

Blodgett, D., Emily, R., Jessica, L., Tad, S., Dwane, Y., (2016). An Analysis of Water Data Systems to Inform the Open Water Data Initiative. (JAWRA) 52(4):845–858.

Briar, D., Cajina, L., Clark, E., Eicher, K., McEnery, J., Perryman, M., Smith, M., and Wilson, G., (2015) IWRSS Interoperability and Data Synchronization: Initial Operating Capabilities - Conceptual Design Report.

Horsburgh, J., Aufdenkampe, A., Mayorga, E., Lehnert, K., Hsu, L., Song, L., et al., (2016). Observations Data Model 2: A community information model for spatially discrete Earth observations. Environmental Modelling & Software, 79, 55-74.

Integrated Water Resources Science and Services - System Interoperability and Data Synchronization: Scoping and Requirements Report (2013) NOAA, U.S. Army Corps of Engineers, U.S. Geological Survey.

McEnery J, Ingram J, Duan Q, Adams T, Anderson L (2005) NOAA's advanced hydrologic prediction service: Building pathways for better science in water forecasting. Bulletin of the American Meteorological Society 86:375–385. doi:10.1175/BAMS-86-3-375

McKay, L.,Bondelid, T., Dewald, T., Johnston, J., Moore, R., andRea, A.,"NHDPlus Version 2: User Guide", 2012

National Oceanic and Atmospheric Administration, (2009) Integrated Water Resources Science and Services, (IWRSS) An Integrated and Adaptive Roadmap for Operational Implementation: Cline, D., compiler, Report Reference #IWRSS-2009-03-02

Pothina, D., & Wilson, A. (2011) Using Python, Partnerships, Standards and Web Services to provide Water Data for Texans. Proc of 10th Python in Science Conference (SCIPY 2011) pp39-42.

Rew, R., Hartnett, E., & Caron, J. (2006). NetCDF-4: Software implementing an enhanced data model for the geosciences. In 22nd International Conference on Interactive Information Processing Systems for Meteorology, Oceanograph, and Hydrology.

Zaslavsky, I., Valentine, D., Whiteaker, T., (2007). CUAHSI WaterML. Open Geospatial Consortium Discussion Paper OGC 07-041r1.