Final Report: A Data Retrieval, Analysis, and Visualization Tool for Hydrological Calibrations and Applications at NOAA National Weather Service

University PI: Yao Liang
Advanced Research Institute
Dept. of Electrical and Computer Engineering
Virginia Tech
4300 Wilson Blvd., Suite 750
Arlington, VA 22203
Email: yaliang@vt.edu

NWS: Thomas Adams at Ohio RFC

Project period: 01/01/04 – 05/31/05
Award: $9,064

Table of Contents

Chapter 1 . Introduction

1.1. Objectives, rationale, approach, significance, and expected benefits

The advent of global observing systems (e.g., satellite remote sensing) and global field programs has generated unprecedented amounts of critical multi-variate data for scientific studies. The advancement of such new kinds of data is creating new opportunities to explore new findings to improve hydrologic calibrations, weather prediction skills, and understanding of critical environmental interactions. In particular, modern information technology (IT), such as data visualization, and data mining, may significantly improve our hydrological and weather predictions and calibration skills through easier access to various data sources.

Objectives:

In this project, we focused on developing a web-enabled XML-based data retrieval, analysis, and visualization tool (HIDE – Hydrological Integrated Data Environment) to facilitate easier access of various data sources with heterogeneous data formats needed for improving hydrological model calibration processes and other hydrological research and application activities at the NWS. The objectives for this project are:

(1). Coherent management, share and visualization of various heterogeneous data sources with dramatically different formats and structures with extensibility, scalability, uniformity, and transparency. The data tool allows full or partial direct access, retrieval, display, and analysis of the NWS River Forecast System (NWSRFS) IHFS database. An easy direct access and retrieval of the OFS data source could significantly facilitate the analysis of the model calibration and post event inspection, filling possible data gaps, and obtaining reservoir operation information for hydrologic forecasts at NWS.

(2). Rapid search and access of massive amounts of data by specifying query conditions, browsing, analyzing, aggregating, and visualization of queried data and customizing the formats of retrieved data, by geographically distributed researchers effectively. Also, the retrieved data could be organized in time-series for individual locations from data to facilitate future distributed hydrological modeling and their calibrations for improvement of hydrologic forecasts at interior locations.

(3). Analysis and visualization of the accessed data first before retrieving them which could facilitate the data selecting process.

(4). Platform Independency with a consistent and unified user-friend interface, based on any type of web browser, such as Netscape or Internet Explorer.

Rationale:

Operational research plays a critical role in the operationally hydrologic forecast mission at the NOAA National Weather Service (NOAA NWS) River Forecast Centers (RFCs) which is uniquely mandated amongst Federal agencies to provide forecasts for the Nation's rivers by providing daily and other forecasts at over 4,000 points across the contiguous United States and Alaska. The operational component of the mission is performed at 13 River Forecast Centers and approximately 120 Weather Forecast Offices at strategic locations across the United States. The NWS Office of Hydrologic Development supports the operational mission by developing, implementing, and maintaining hydrologic models and systems. The forecast models used are developed and calibrated for specific rivers and streams based on historical events. They are conditioned and constrained operationally using *current* observations and, in the case of operational ensemble forecasts, with historical data as well. Therefore, delayed, inaccurate, inconsistent, incomplete or insufficient data used for calibrating the forecast models can cause significant problems in the forecast process and for the forecasters who operate them. In order to make it easier to access different data sources needed for improving the hydrological model calibration processes for achieving better forecast skills, significant efforts have been made to develop a broad set of tools over the years at the Hydrology Laboratory of the NWS. However, these developed tools are limited for specific data sources at present. As new data sources emerge and research with new models progresses, corresponding tools for exploring, accessing, analyzing these new heterogeneous data sources have not been developed. Clearly, if such challenges and issues in data acquisition, processing, and management are not sufficiently addressed, significant improvement in hydrologic research cannot be achieved.

One of the possible solutions to the above mentioned problem is to develop data integration models in the information system super-imposed on the specific scientific domain. This methodology works well as long as the models can satisfy the constraints and rules of the domain, with the advent of ontology to characterize the semantic nature of the data. However, the relative autonomy nature of the data collecting organizations, e.g. US Geological survey [5] and the presence of legacy systems prove to be a hindrance in achieving complete semantic interoperability. Our approach, an integration solution achieves semantic and structural interoperability through a generic concept "DataNodes" arranged as a DataNode tree. The generic nature of the DataNodes can very well be applied in the context of the semantic heterogeneity and structural heterogeneity of the registered data sources. The implementation of our DataNode tree model relies on a metadata model for data semantics and logic organization, which eliminates/reduces the collaborative effort required from the participating data sources. The approach is applied onto the datasets from USGS and NWSRFS IHFS database.

Benefits:

This work directly addresses two key concerns in data management and analysis: (1) new data analysis techniques with wide application and usefulness in teaching, research, or operational forecasting, and (2) facilitate improvement of hydrological forecast through improvement of hydrological model calibration. The tool developed as part of this project could significantly eliminate current burdens and efforts of accessing massive data of diverse data sources (e.g., data searching, selecting, retrieving, aggregating, preprocessing, and analyzing). Therefore, the data tool could offer great potentials to improve operational hydrological forecasts by facilitating the hydrological calibration processes and research. Also, the data tool has great potential to benefit other RFCs through a future plan of national implementation of its successful components. The features supported by the data tool are:

- Access and manage very large volumes of heterogeneous data from distributed diverse sources with distinct different structures and formats promptly, intelligently, and efficiently,
- Easily adopt changes, over times, of transmitted data structures and formats due to rapid advancement in data transmission technologies (i.e., open system architecture),
- Handle multi-resolution data outputs and provide various customizable data output formats,
- Provide users with easy and uniform GUI (Graphical User Interface) type of data access across heterogeneous data sources without worrying about computer platform and different data formats and structures used in any data sources,
- Merge datasets obtained from diverse data sources as needed, and provide various data analysis methods, and
- Provide various 2-D and 3-D visualization methods to visualize data from heterogeneous data sources and modeling results.

Chapter 2 . HIDE: Mechanisms, Architecture and Implementation

The complexities in data integration are attributed to various heterogeneities of the data as well as the data systems. One can define several kinds of heterogeneity leading to several levels of interoperability: system, syntax, structure and semantic [4]. In this classification, the machine readable aspects of data representation falls into syntactic heterogeneity, and the representational heterogeneity in terms of data modeling falls into structural heterogeneity. Semantic heterogeneity relates to the difference in the semantics of the datasets to be relevant to semantic interoperability. The semantic interoperability requires that system understands the semantics of the information request as well as those of the information sources and satisfy the request as much as it can. The work in [2], list some of the key issues to be considered in a scientific data integration scenario: the nature of the datasets (structure, schema, syntax etc), publishing methodology which is often non-standardized resulting in an "annotation pipeline", and the fluid and dynamic federation of the datasources.

Current research trends in a scientific domain are exploring the possibilities of using ontologies for semantic interoperability. The success of these efforts depends on a comprehensive ontology design complete with domain specific generic rules. Ontology is an explicit specification of conceptualization [3]. Consequently, an ontology defines a set of terms of the domain (e.g.: classes, objects, relations, properties) and formal axioms which constrain the interpretation and well formed use of the terms. Defining ontologies helps systems (agents) in communication and collaboration, to bridge the semantic gap between information systems in a scientific domain.

The integration architecture of HIDE utilizes partial collaborative effort from autonomous data sources. A user can define various characteristics such as data organization, structure, and semantics of the datasets to be integrated through the metadata standard. Our approach achieves a balance between the requirements of autonomous control over data by hydrologic data sources and a uniform infrastructure satisfying semantic needs. The uniqueness of our approach is to combine the ontology aspects of the domain and the logical data organization exclusive to a data source through a tree model "DataNode Trees". Consequently this helps in attaining a "virtual information space" uniform to all data sources. The tree model provides the flexibility of easy query evaluations and representations. Furthermore, the model can be specified using our metadata standards.

2.1. Methods and Mechanisms

Data Integration Model

In our Data integration model we integrate heterogeneous data by defining a generic concept "DataNode". A DataNode, the smallest unit of information integration can represent an ontological concept such as precipitation or a structural elements such as database, tables, files, datasource etc. Based on the temporal-spatial nature of hydrology scientific data, a DataNode in the hydrology datasets can be modeled as a Time-Space-Attribute node. In other words, a DataNode essentially represents Time-Space-attribute information. The Time aspect of the model corresponds to factors for instance; realtime, monthly, daily while the space aspect characterizes spatial features such as states, watersheds, latitude-longitude ranges. The "attribute" aspect of the model facilitates in defining special variables/features that often qualifies a unit of hydrologic information. . For example, in case of modelling precipitation data, the variable "precipitation" can be considered as an attribute. Similarly while defining a unit of hydrologic information such as a data source USGS, the feature "USGS" can be expressed as an attribute.

The association between various DataNodes in our system conceptualizes a generic ontology of the domain (*domain view)* or a logical organization of data (*user view).* The *views* in the system define relations between the DataNodes as hierarchical and are presented to the user as a DataNode sub trees. These views are later combined using *Adaptation DataNodes* into a DataNode tree (Figure 1). Generalization of semantic and structural nuances in a datasource through DataNodes and corresponding DataNode tree presents transparency to user. Hence, user can pose high level queries independent of the "DataNode tree" which are translated into DataNode sub-queries and executed.

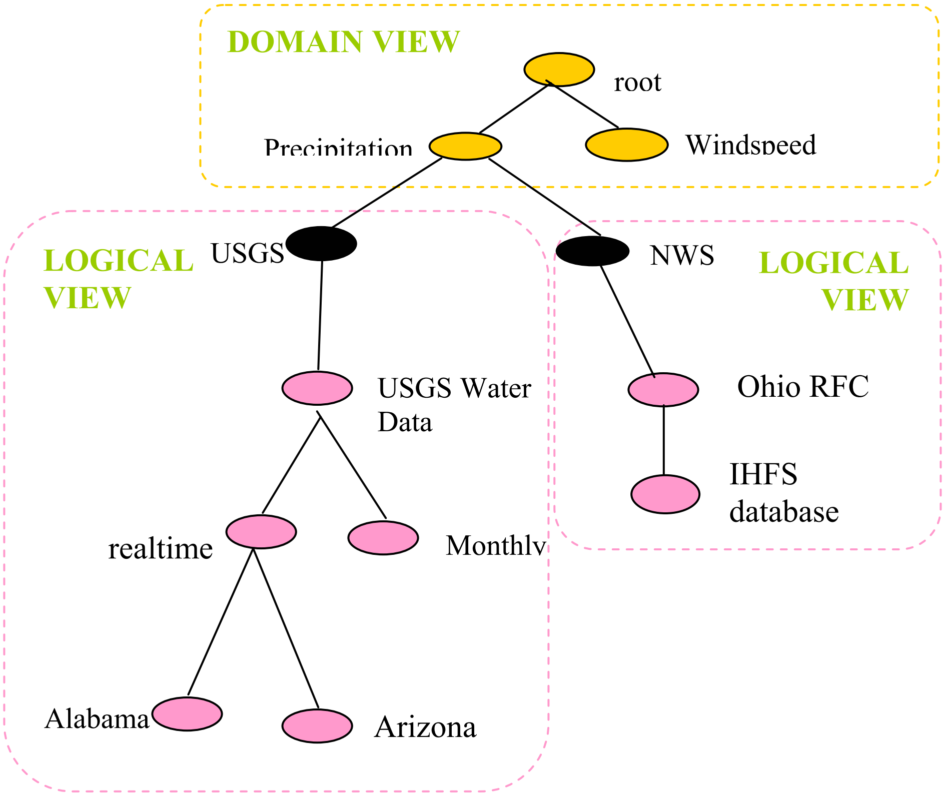


Figure 1. Views – Domain View and logical views of data sources USGS and NWS.

The *domain view* of a DataNode tree can be defined for various users wit different levels of permissions while the *logical views* is unique to a datasource.

Search and Query a dataset is translated to "search" and "query" DataNodes of the DataNode trees in our system. Based on the user defined parameters, our search engines traces the search from root DataNode to any intermediate or leaf node (Figure 2). If the found DataNode is not satisfactory to the user, the system provides a guided search by taking the user to each levels of the DataNode tree. Once the DataNodes are properly identified, the users can pose "query" to the query interfaces of the DataNode. The DataNode query at the leaf node referred as "atomic query" is off special interest as it performs the actual query at the datasource. The DataNode query at the intermediate levels of the logical view is an "aggregation query" involving, compose of multiple sub queries and delegating it to its immediate children. The child DataNodes applies the sub-DataNode query with the continuation of the procedure until the leaf node. At the leaf node, a DataNode "aggregation query" is transformed to an "atomic query". The results of all sub-DataNode queries are aggregated/joined at the intermediate DataNodes.

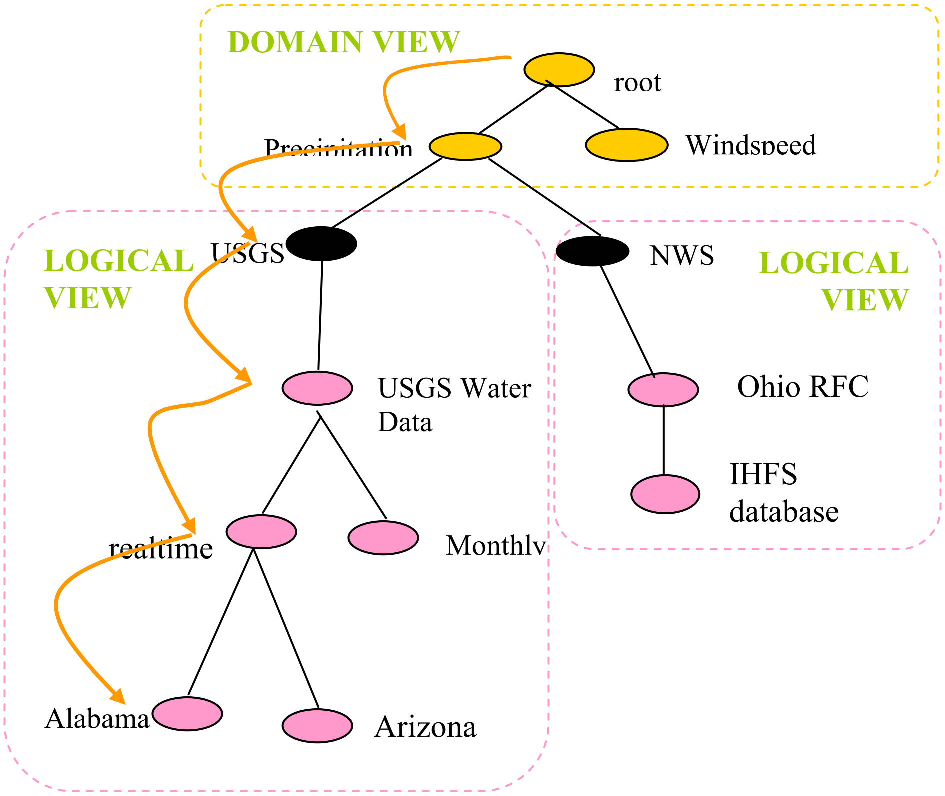


Figure 2: A query Trace for the query "“What is the precipitation distribution of realtime water data for the state Alabama from USGS data” source”.

Metadata Standards

Keeping the design objective of flexibility and extensibility, we have used the concept of metadata for defining our DataNode trees, the query interfaces and other necessary information. The XML language is used to describe the metadata. Our XML metadata standard can be classified into 4 categories.

1. Describe the DataNode & DataNode tree.
2. Define query model and its translation to the query model of the datasource
3. Define the syntactic nature of data.
4. Define data transformations if required.

The first standard is used to define a DataNode and DataNode tree. In this specification, user needs to include information about the DataNode such as name, documentation (URLs, weblinks), vocabulary (a collection of similar names), links to specification of its children and other additional info. Disparity of interfaces of each data sources makes locating and access to data cumbersome. Hence one of our primary objectives is to provide a uniform representation of the user interfaces irrespective of the complexities of underlying data source. This is achieved through our 2nd standard. In this specification, the user can define a query interface for each DataNode. In addition, each datasource uses various methodologies for storing and accessing their data; for instance websites, databases, files/directories. Hence, we use a translation mechanism, which transforms the query posed to the DataNode, to the query to the data source. The 3rd standard is used to define the syntactic details of the data such as ascii files, comma-separated etc. To facilitate the need for a common data model for different data, our system implements 3 types of data models – temporal, temporal-spatial and spatial model. Our final standard can be used to transform the data to these models. A sample of our specification is shown in Figure 3.

```
<?xml version="1.0"?>

<DataNode xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation='c:\IDAM\xml\xsd\funcModelSchema.xsd' >
        <identifiedAs>
                <name> River Forecasting </name>
                <label> river forecasting </label>
                <index> rfc,river,forecasting,center </index>
        </identifiedAs>
        <documentedBy>
                <url> www.noaa.gov </url>
                <history> River forecast data from sources such as Ohio River forecast
center</history>
        </documentedBy>
        <DataNodeMembers>
                <DataNodeMember>
                        comet/ohio/ohiorfc_funcModel.xdms
                 </DataNodeMember>
        </DataNodeMembers>
</DataNode>
```

Figure 3: The Description specification (1st standard) for the DataNode "Ohio RFC" in Figure 2.

2.2. Architecture

There are various efforts in designing strategies to address the extensibility of the system. One such approach is to use a modular architecture in which a system is composed of small and autonomous components that collaborate to achieve a common goal. The beauty of this architecture is that one can add/replace any components without affecting others. Based on this aspect, it fit perfectly well to the nature of the HIDE system (Figure 4).

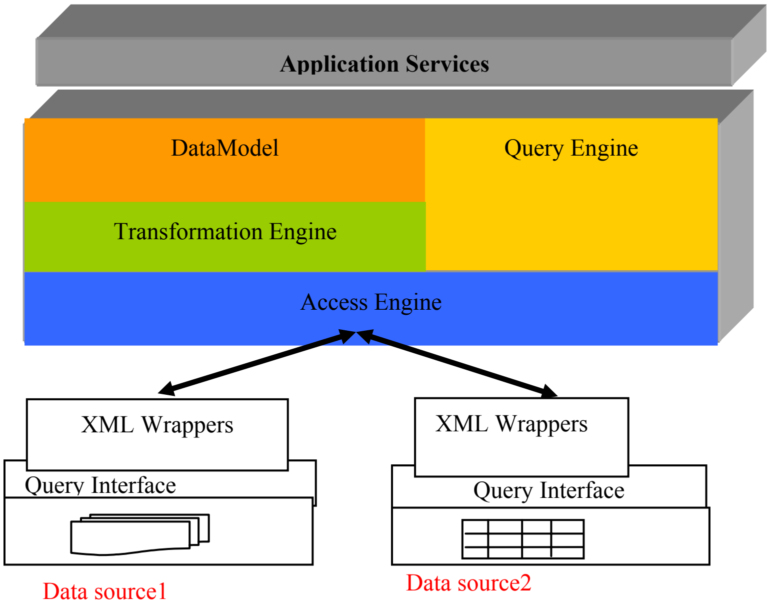


Figure 4: The modular architecture of HIDE.

Application services

A hydrologist is often required to perform various kinds of data analysis on the data. Some of the analysis tools used are visualization (2D, 2D-3D transform), search engines, forecasting tools and modelers. The application services module of HIDE is used to present these services. As the interface between each module is open, new application services can be identified and plugged in without significant changes in the subsequent modules

Data Models

The interim DataModels of HIDE- temporal, spatial and temporal-spatial facilitates a common representation of heterogeneous data in the system. The definition of these models is based on the nature of the hydrology data. The model defines a uniform representation of the data while incorporating the data integration aspects as well. For instance; precipitation data from USGS are characterized differently compared to precipitation data from NWS. Addressing this requirement, the DataModels of HIDE are linked to the corresponding DataNodes. Additionally, one can say, the DataNode tree in the system acts as core engine driving the various modules.

Query Engine

One of the key process in data integration is the evaluation of query from the user to the data sources. The query engine module, as the name suggests, acts as the query processor of the system. It transforms the user level query to intermediate query and performs search and query operation. The query evaluation is a complicated process that involves traversal of the tree and can often become a bottleneck in the system. Therefore, successful completion of the query evaluation depends on the optimized creation of the DataNode tree from the metadata.

Transformation Engine

The complex nature of scientific datasets can be defined using various standards such as DODS [2004]. Each of these standards has a distinct data model. For example; DODS has defined data types of ‘Structure’, ‘Sequence’, ’Arrays’ and ‘Grids’ for modeling various kinds of scientific data.

Although the basic nature of the data is temporal-spatial, a certain extent of transformation is required for the data from the source to the interim DataModels of HIDE. The transformation engine performs this transformation.

Access Engine

Access Engine module performs the transformation of the intermediate query of the HIDE system to query of the external data sources. In addition to that, the transformed query is posed to the data source with the retrieval of the data.

2.3. Implementation

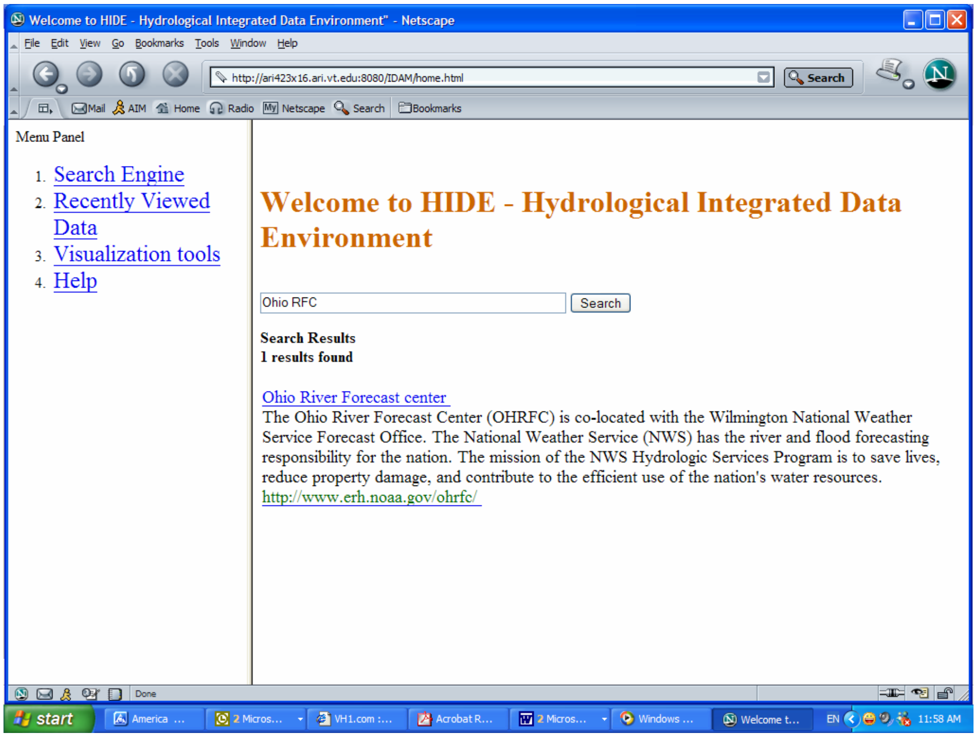
We have implemented our system on a Java environment. Our system is web-enabled and uses a client-server approach. We use Apache-Tomcat server running at back-end. By using a web-based method with a Java runtime environment, we were able to address the objective of platform independency in our system. Some of the features provided in the system are search and query, visualization tools, history of retrieved data, saving the data to local file systems.

In the current version of our system, two types of heterogeneous data from datasources USGS and NWSRFS OFS are integrated. We have also provided the functionality of saving the data to OH datacard format.

The features of the system are built and supported by the model DataNode trees. Some of them are explained below.

Search Engine: We have developed a Search Engine similar to “Google” which can be used to find the datasets. The primary functionality of the search engine is to perform a search to find appropriate datasets based on the search string provided by the user. The search is carried out on the model DataNode trees, traversing from the root node to the appropriate nodes corresponding to the search string. The user can view a list of DataNodes supplemented with sample information after the search is performed. Upon selecting a DataNode, the user can perform a query through the query interface provided.

The Search screen after performing a search for the Ohio RFC dataset is shown below.


Welcome to HIDE - Hydrological Integrated Data Environment" - Netscape
File Edit View Go Bookmarks Tools Window Help
http://ari423x16.ari.vt.edu:8080/IDAM/home.html
Search
Mail AIM Home Radio Netscape Search Bookmarks
Done
start America ... 2 Micros... VH1.com :... Acrobat R... 2 Micros... Windows ... Welcome t...
EN 11:58 AM

Menu Panel

1. Search Engine
2. Recently Viewed Data
3. Visualization tools
4. Help

Welcome to HIDE - Hydrological Integrated Data Environment

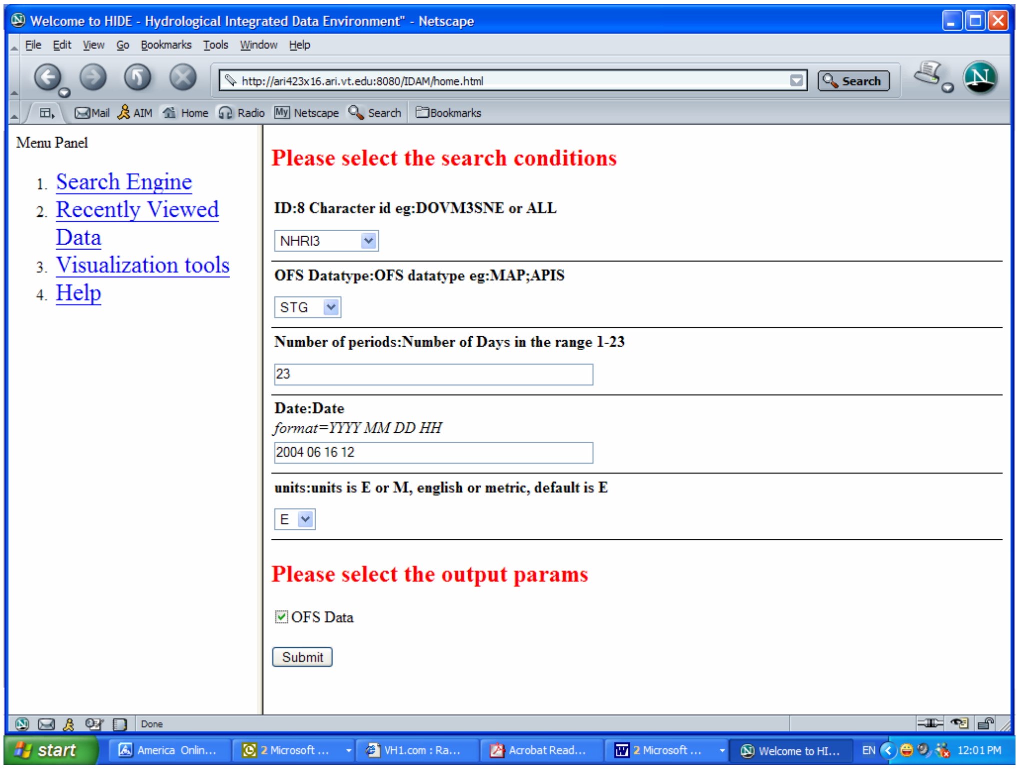
Ohio RFC [Search]

Search Results
1 results found

Ohio River Forecast center
The Ohio River Forecast Center (OHRFC) is co-located with the Wilmington National Weather Service Forecast Office. The National Weather Service (NWS) has the river and flood forecasting responsibility for the nation. The mission of the NWS Hydrologic Services Program is to save lives, reduce property damage, and contribute to the efficient use of the nation's water resources.
http://www.erh.noaa.gov/ohrfc/

Figure 5: Search Engine.

Upon selecting the DataNode, the user can view the query interface for the node. These interfaces are developed based on the XML based configuration query metadata provided by the System integrator. In the metadata, user can include various search conditions and possible output parameters. It has to be noted here that the modifications to the system if a different storage mechanisms for instance “database”, used will be minimal. Hence the query interfaces remains the same with the previous versions.


Welcome to HIDE - Hydrological Integrated Data Environment" - Netscape
File Edit View Go Bookmarks Tools Window Help
http://ari423x16.ari.vt.edu:8080/IDAM/home.html
Search
Mail AIM Home Radio Netscape Search Bookmarks
Done
start America Onlin... 2 Microsoft ... VH1.com : Ra... Acrobat Read... 2 Microsoft ... Welcome to HI...
EN 12:01 PM

Menu Panel

1. Search Engine
2. Recently Viewed Data
3. Visualization tools
4. Help

Please select the search conditions

ID:8 Character id eg:DOVM3SNE or ALL

NHRI3

OFS Datatype:OFS datatype eg:MAP;APIS

STG

Number of periods:Number of Days in the range 1-23

23

Date:Date
format=YYYY MM DD HH

2004 06 16 12

units:units is E or M, english or metric, default is E

E

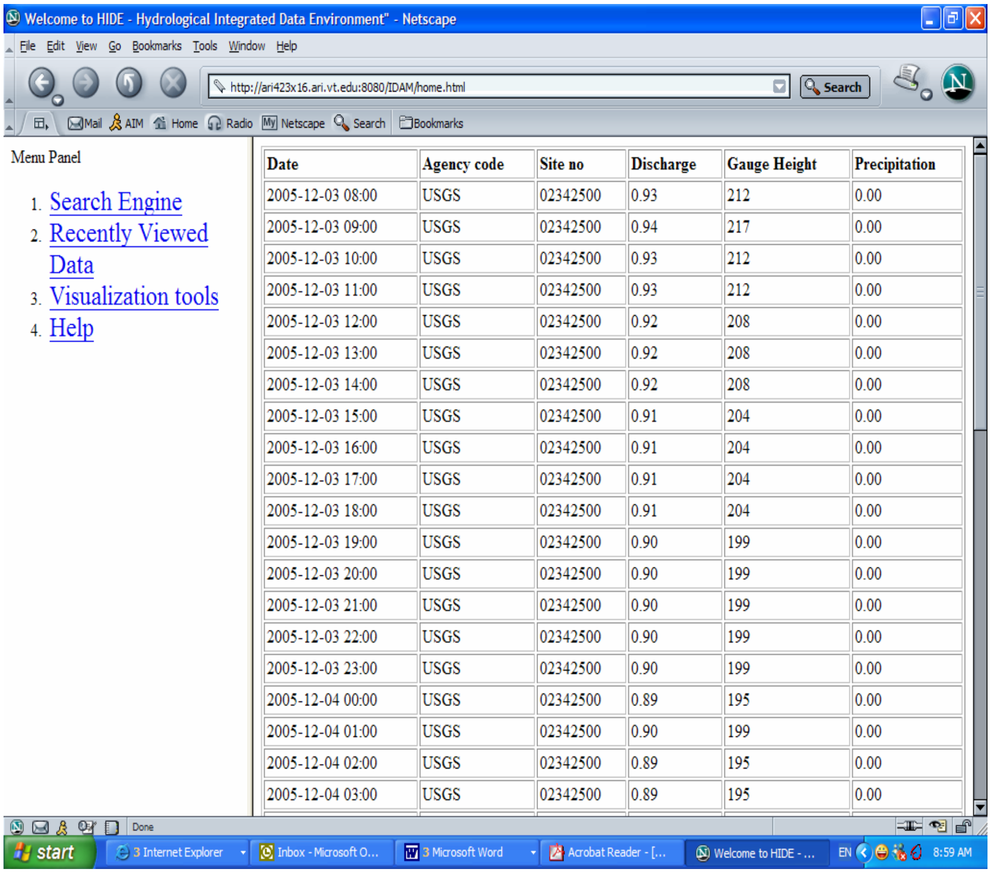
Please select the output params

☑ OFS Data

[Submit]

Figure 6: Query Interface.

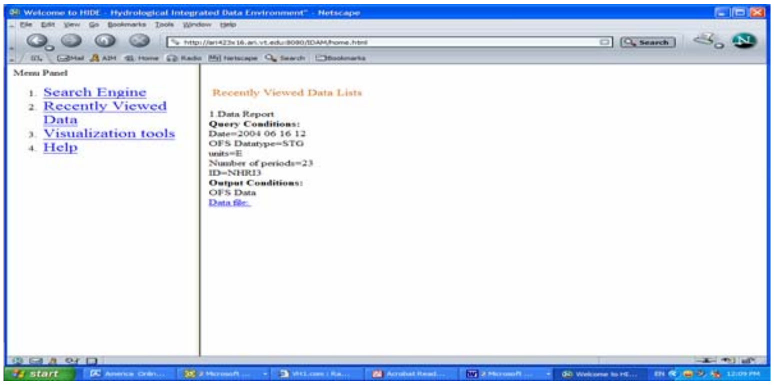
Data Views: Upon entering the search conditions and possible output parameters in the query interface, HIDE sends the query to the data engine(s), retrieve the results, performs necessary transformations and integration for multiple data engines and results are displayed. Though the present version, HIDE system has the capabilities to issue SQL queries to any database, it does not have necessary tables to show the forecasted and current data (*waiting for inputs from Ohio RFC*). Hence the DataViews of the USGS data is shown below (Figure 7).


Welcome to HIDE - Hydrological Integrated Data Environment" - Netscape
File Edit View Go Bookmarks Tools Window Help
http://ari423x16.ari.vt.edu:8080/IDAM/home.html
Search
Mail AIM Home Radio Netscape Search Bookmarks
Menu Panel
1. Search Engine
2. Recently Viewed Data
3. Visualization tools
4. Help
Done
start
3 Internet Explorer
Inbox - Microsoft O...
3 Microsoft Word
Acrobat Reader - [...
Welcome to HIDE - ...
EN
8:59 AM

Date	Agency code	Site no	Discharge	Gauge Height	Precipitation
2005-12-03 08:00	USGS	02342500	0.93	212	0.00
2005-12-03 09:00	USGS	02342500	0.94	217	0.00
2005-12-03 10:00	USGS	02342500	0.93	212	0.00
2005-12-03 11:00	USGS	02342500	0.93	212	0.00
2005-12-03 12:00	USGS	02342500	0.92	208	0.00
2005-12-03 13:00	USGS	02342500	0.92	208	0.00
2005-12-03 14:00	USGS	02342500	0.92	208	0.00
2005-12-03 15:00	USGS	02342500	0.91	204	0.00
2005-12-03 16:00	USGS	02342500	0.91	204	0.00
2005-12-03 17:00	USGS	02342500	0.91	204	0.00
2005-12-03 18:00	USGS	02342500	0.91	204	0.00
2005-12-03 19:00	USGS	02342500	0.90	199	0.00
2005-12-03 20:00	USGS	02342500	0.90	199	0.00
2005-12-03 21:00	USGS	02342500	0.90	199	0.00
2005-12-03 22:00	USGS	02342500	0.90	199	0.00
2005-12-03 23:00	USGS	02342500	0.90	199	0.00
2005-12-04 00:00	USGS	02342500	0.89	195	0.00
2005-12-04 01:00	USGS	02342500	0.90	199	0.00
2005-12-04 02:00	USGS	02342500	0.89	195	0.00
2005-12-04 03:00	USGS	02342500	0.89	195	0.00

Figure 7: DataViews of USGS data.

Recently Viewed Data: This functionality allows the user to keep track of all the datasets which he/she has recently used. The following screen shows the recently viewed datasets.

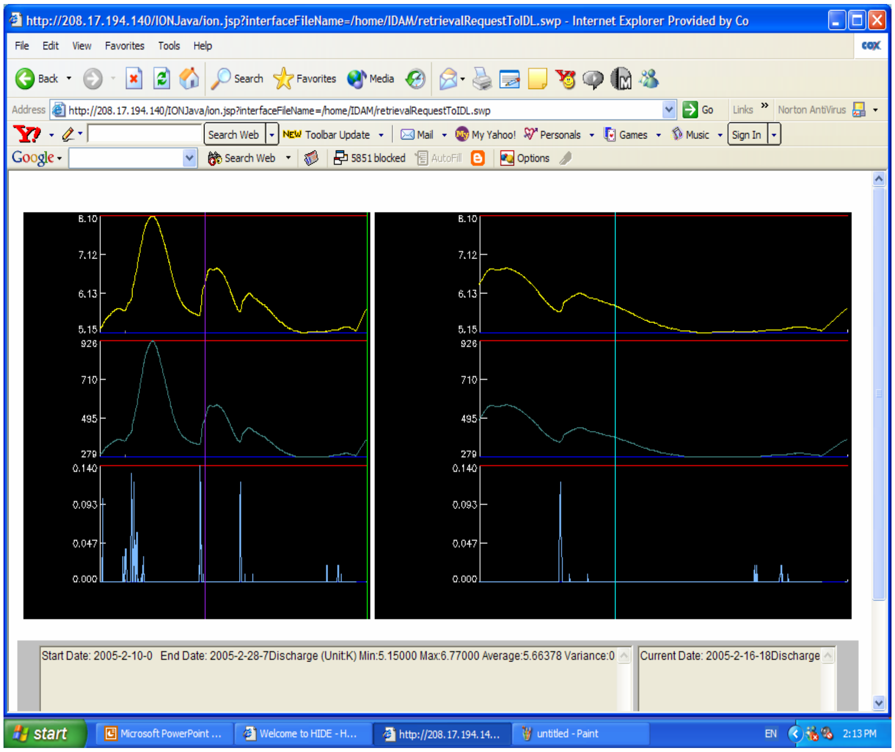

Welcome to HIDE - Hydrological Integrated Data Environment" - Netscape
File Edit View Go Bookmarks Tools Window Help
http://ari423x16.ari.vt.edu:8080/IDAM/home.html
Search
Mail AIM Home Radio Netscape Search Bookmarks
Menu Panel
1. Search Engine
2. Recently Viewed Data
3. Visualization tools
4. Help
start
12:09 PM

Recently Viewed Data Lists

1.Data Report
Query Conditions:
Date=2004 06 16 12
OFS Datatype=STG
units=E
Number of periods=23
ID=NHRI3
Output Conditions:
OFS Data
Data file:

Figure 8: Recently used Data.

Visualization tools: This functionality allows the user to apply our remote visualization tools on the data to be retrieved. The tools provided in the system for temporal data are 2D plots (Figure 9a), and spatial data are 3D plots (Figure 9b).

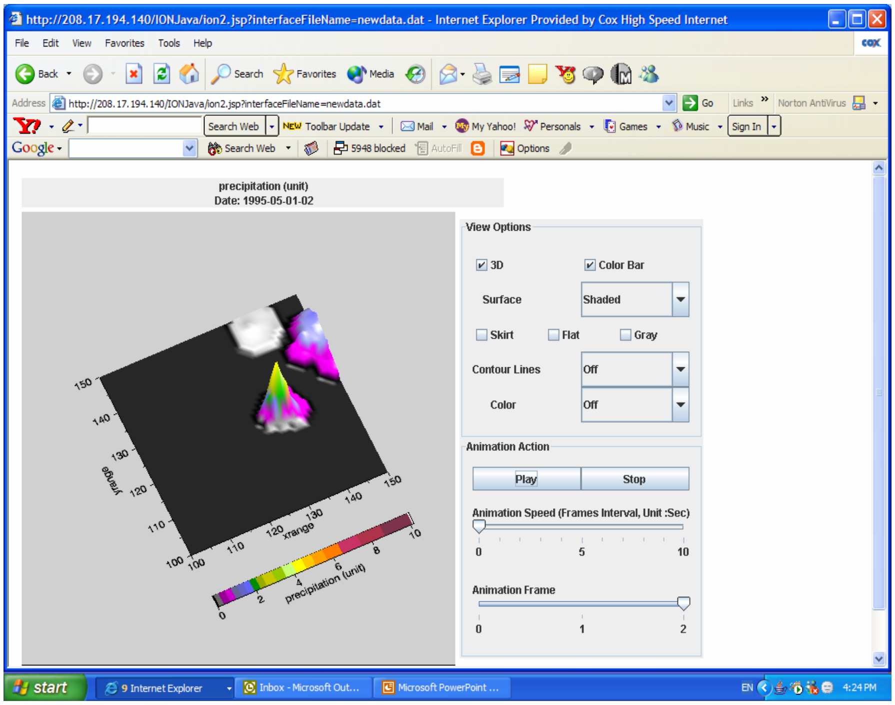


http://208.17.194.140/IONJava/ion.jsp?interfaceFileName=/home/IDAM/retrievalRequestToIDL.swp - Internet Explorer Provided by Co

Start Date: 2005-2-10-0 End Date: 2005-2-28-7Discharge (Unit:K) Min:5.15000 Max:6.77000 Average:5.66378 Variance:0

Current Date: 2005-2-16-18Discharge

Figure 9 (a): Visualization tool- 2D plot. The Left window shows the temporal plots of parameters such as precipitation etc. The Right window (zoom window) shows the zoomed range (purple – green) of the Left window. The Mean, STD, Minimum for each parameter is also shown.



http://208.17.194.140/IONJava/ion2.jsp?interfaceFileName=newdata.dat - Internet Explorer Provided by Cox High Speed Internet

precipitation (unit)
Date: 1995-05-01-02

View Options

☑ 3D ☑ Color Bar

Surface: Shaded

☐ Skirt ☐ Flat ☐ Gray

Contour Lines: Off

Color: Off

Animation Action

Figure 9 (b): Visualization tool- 3D plot for precipitation data. The right window shows various options such as animation, color bar, grey scaling that can be applied to the plot.

Chapter 3 . Accomplishments

The focus of our project was to develop a web-enabled XML-based data integration, retrieval, analysis and visualization tool. This tool will enable hydrology scientists at NWS to access, analyze and share distributed heterogeneous massive amounts of data and NWS's own data more resourcefully and efficiently. The accomplishments of this project are summarized as follows:

1. A Data integration model for integration of data with heterogeneities – semantic, structure, system and syntax from autonomous data sources based on DataNode trees. The model is generic to be applied to other domain as well.

2. The implementation of the model – HIDE can be used to integrate various data from internal NWSRFS to external data sources USGS, etc. dynamically. This is necessary due to the fact that only queried data is integrated while the huge dataset is residing at the data source.

3. The layered architecture of HIDE promotes simplicity and extensibility thus facilitating the addition of new application services, new data models easier and simpler.

4. An XML-based implementation of the system helps in addressing the objective of flexibility. Hence addition of a new DataNode/ Dataset to the system involves the mere inclusion of new metadata files. This helps tremendously as the system need not be modified for every data set changes.

5. The implementation of the remote data visualization tools for both 2-D and 3-D data online visualization.

6. Additional enhancements such as SQL capability facilitating the data integration for data in databases. Therefore, data in any data storage mechanisms such as databases, files, websites, etc., can be integrated by HIDE. We extended the HIDE system so that it will be able to integrate the data from IHFS database and provided all the functionalities of the HIDE system such as search, query, data views, and visualizations. However, due to the dramatic change in the data storage and management platform of the NWS Ohio RFC during our project period, VT team has not been able to receive the NWS' database schema and data of requisite tables (current and forecasted data) yet, and therefore the data from the IHFS database have not been actually integrated into HIDE yet. We expect to receive the related data schemas and data from Ohio RFC in the near future, so that we can integrate Ohio RFC's actual data and test on it.

References

[1] S.S Anand and A.G Büchner, “Decision Support Using Data Mining”, Financial Times Management, ISBN 0-273-63269-8, pp. -168, London: Pitman Publishers, 1998.

[2] O. Boucelma et al., “Report on the EDBT'02 panel on scientific data integration,” ACM SIGMOD Record, vol. 31, no. 4, Dec. 2002, pp. 107-112.

[3] T. Gruber, “Toward principles for the design of ontologies used for knowledge sharing,” Int’l J. Human-Computer Studies, vol. 43, no. 5-6, Dec. 1995, pp. 907-928.

[4] A. Sheth,”Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics,” Interoperating Geographic Information Systems, M.F. GoodChild et al., eds., Kluwer Publishers, 1998.

[5] USGS, U.S Geological Survey, 2004; http://waterdata.usgs.gov/nwis.