CUAHSI

universities allied for water research

CUAHSI HIS CENTRAL 1.2

Web based Data Service Repository for HIS Web services

April 4, 2010

by:

Tom Whitenack
San Diego Supercomputer Center
University of California, San Diego

Distribution

Funding

This document was developed as part of the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) Hydrologic Information System project, under NSF Grant No. EAR-0413265. Much input and feedback has been received from the CUAHSI Hydrologic Information System development team. Portions of this document borrow from the "HIS Central 1.1 Display and Usability Functional Specifications (10/3/2008), and the earlier "HISCentral 1.1" (3/31/2009) description. Contribution of team members is acknowledged here

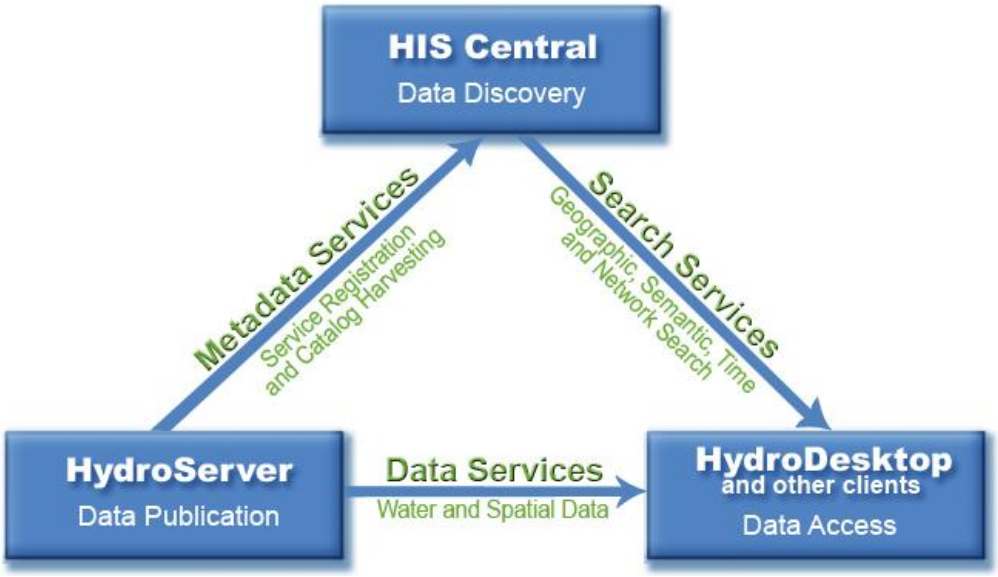
TABLE OF CONTENTS

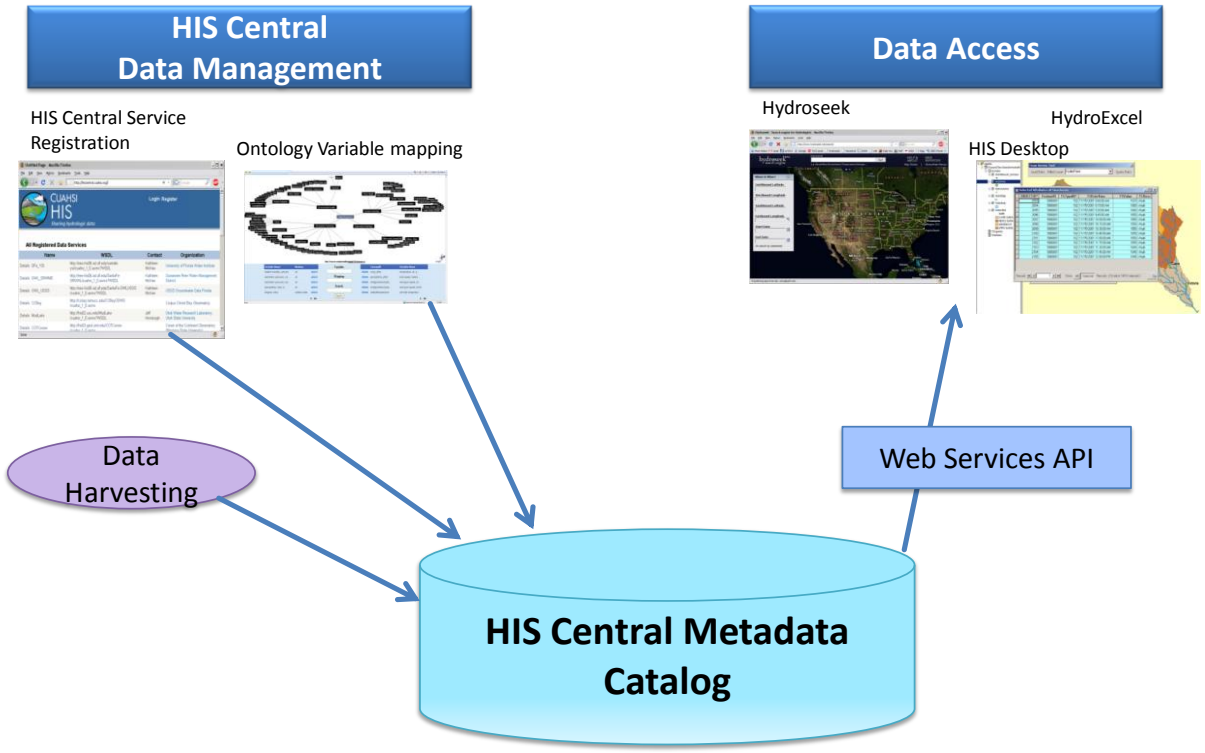
INTRODUCTION

HIS Central is the web application which provides an interface for adding and managing registered water data services and the HIS central metadata catalog. The central metadata catalog is designed to maintain observation series information, including site information, variable information, period of record, as well as project metadata – for all registered data sources of hydrologic observations. These two components are in the center of the CUAHSI HIS architecture.

Several user roles are associated with HIS Central functionality. *Data managers* who have gone through the task of creating a WaterOneFlow data service for their observational data (possibly using HISServer or Hydroserver), use it to register the service to the central registry, to make it possible for HIS users to discover and access the data, and integrate it with information from other similarly registered data services. *HIS Central Administrator* is responsible for maintaining hydrologic metadata from sources other than those compatible with WaterOneFlow services, i.e. federal agency repositories. In addition, this role is responsible for periodically updating the metadata catalog and managing all aspects of the HIS Central application. The metadata catalog also maintains a hydrologic ontology so that well established concepts can be used to search across multiple water data services which may be using different terminology for their variables. An *ontology curator* is responsible for maintaining and evolving the ontology, and managing associations between ontology concepts and variables. Once the data are published and registered in HISCentral, they are accessed by *metadata catalog users*, via multiple online and desktop clients. In addition to the concept-based search enabled by the ontology mappings, the metadata catalog maintains spatial and temporal indexes assisting in spatio-temporal searches across multiple observation series.

The place of HISCentral and the central metadata catalog in HIS is demonstrated in the triangular diagram of HIS architecture, and a follow –up more detailed schema.





CONTENT OF THE DOCUMENT, AND CHANGES COMPARED TO HISCENTRAL 1.1

This document describes version 1.2 of HISCentral, available as of April 2010. Compared to HISCentral 1.1 developed in 2009, the new application includes improved user interface, incorporates *ontology tables*, and supports the new ontology management approach through all steps of metadata processing: from publication and variable tagging to data discovery. Another major addition is *HISCentral web services* that expose the content of the metadata catalog and the ontology in a way that is convenient for client applications such as HydroDesktop and HydroViewer. In addition to the overall HISCentral description, the document describes these additional components and provides an in-depth review of the internal structure of the HIS Central metadata catalog and catalog update approaches that have been implemented to date.

Related HIS specification documents, which contain additional information about some aspects of HISCentral functionality, are available on the CUAHSI HIS web site (his.cuahsi.org) and include:

- description of Hydrodesktop application (the part where it the application accesses the metadata catalog is related);
- specification of the Observations Data Model (with which HISCentral catalog shares semantics and – to a significant degree – table structure);
- specification of WaterML (with which HISCentral services share many output objects), and

- functional descriptions of HydroSeek and Hydrotagger (which are incorporated in the HISCentral applications with functionality largely unchanged but completely reworked underlying ontology management system).

METADATA CATALOG

The metadata catalog is the database that maintains metadata (at the level of observation series) for all of the registered water data services. The database schema is similar to ODM, but differs in that it contains additional tables with data services (Networks) and ontology information, and does not contain the DataValues table. The database follows a star schema centered on the SeriesCatalog table. For each service, the database stores a collection of sites, or stations; each site has one or more data series. Each data series is defined by the site where the measurements were taken, what variable is being monitored, when it started, and how many data values there are, as well as several additional metadata characteristics such as sample medium. Each Variable is tagged with a concept ID from the hydrologic ontology. The core tables of the database schema are pictured below, while ontology tables are described in a separate section.

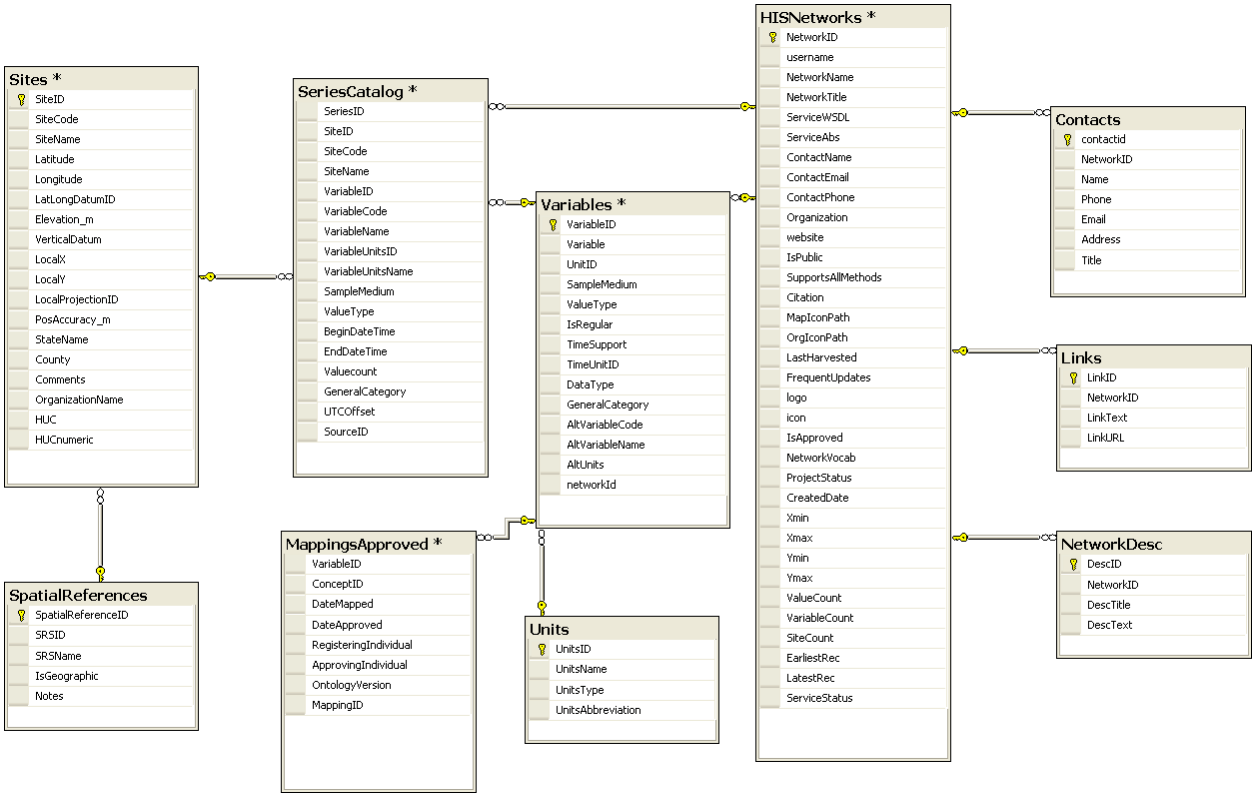


TABLE DESCRIPTIONS:

HISNETWORKS:

This table contains metadata about registered WaterOneFlow services and observation networks they represent.

- NetworkID: Unique database identifier

- Username: logged in user identity
- NetworkName: The name of the network as specified when adding the network:
- NetworkTitle: The human readable name of the network
- ServiceWSDL: The URL end point of the service being registered with "?WSDL" appended
- ServiceAbs: The service abstract.
- ContactName,ContactPhone,ContactEmail: Primary contact information for the service.
- Organization: The organization responsible for data being served
- Website: Organization's website.
- IsPublic: Boolean field indicating whether or not the service is visible to other clients.
- Citation: The citation that is to be included in downloaded data.
- Icon: Image that has been uploaded to serve as a map icon for locations of observation sites available from this service.
- Logo: Image that has been uploaded to represent the source organization.
- IsApproved: Boolean that determines whether or not the service is visible. Similar to the IsPublic but the IsApproved is managed by the HIS Central administrator.
- FrequentUpdates: Boolean indicating if a service is to be included in the weekly harvesting.
- ProjectStatus:
- Xmin, Xmax, Ymin, Ymax: The lat-long bounding box of the sites. These values are calculated during the harvest.
- ValueCount: Total number of values across all available data series. This is summarized from the SeriesCatalog. These values are calculated during harvest.
- VariableCount: Total number of variables, calculated during harvest.
- SiteCount: Total number of sites in the service, calculated at harvest time.
- EarliestRec: DateTime field indicating the earliest available datavalues across all series within a service.
- LatestRec: : DateTime field indicating the lateset available datavalues across all series within a service.

CONTACTS:

The purpose of this table is to allow multiple contacts to be associated with registered service. These contacts are displayed on the public service details page.

- ContactID: unique database identifier
- NetworkID: foreign key to HISNetworks table
- Name, Phone, Email, Address, Title: the contact information

LINKS:

The purpose of this table is to provide on the public details page, any additional links which will provide additional information about the service.

- LinkID: Unique database identifier.
- NetworkID: foreign key to HISNetworks
- LinkText: The text which will appear in the link.
- LinkURL: The URL address the link will take you to.

NETWORKDESC:

The purpose of this table is to provide on the public details page, additional details about the service which may not be included as part of the abstract. Multiple descriptions are possible, allowing the publisher the flexibility to describe their service the way they want. a description of the observation network associated with the service being registered.

- DescID: Unique database identifier.
- NetworkID: foreign key to HISNetworks
- DescTitle: The text which will appear as the header for the description.
- DescText: The body text of the description.

SERIES CATALOG:

Semantics of this table follows ODM specification.

- SeriesID: Unique database row identifier.
- SiteID: Foreign key to sites table.
- SiteCode: as specified in the ODM.
- SiteName: as specified in the ODM.
- VariableID: foreign key to variables table
- VariableCode: as specified in the ODM.
- VariableUnitsID: foreign key to Units table
- VariableUnitsCode: as specified in the ODM.
- SampleMedium: as specified in the ODM.
- ValueType: as specified in the ODM.
- BeginDateTime: as specified in the ODM.
- EndDateTime: as specified in the ODM.
- ValueCount: as specified in the ODM.
- GeneralCategory: as specified in the ODM.
- UTCOffset: as specified in the ODM.
- SourceID: foreign key to HISNetworks table

SITES:

Semantics of this table follows ODM specification.

- SiteID: Unique database identifier
- SiteCode: As specified in ODM
- SiteName: As specidied in ODM
- Latitude: As specidied in ODM
- Longitude: As specidied in ODM
- LatlongDatumID: As specidied in ODM
- Elevation_m: As specidied in ODM
- VerticleDatum: As specidied in ODM
- LocalProjectionID: As specidied in ODM
- OrganizationName: This is the NetworkName and is used to relate site to the HISNetwork table.
- Huc8: Eight digit hydrologic unit code (as a string) which the site falls within.
- Hucnumeric A numerical representation of the HUC code, which is used when searching by watersheds.

VARIABLES:

Semantics of this table follows ODM specification.

- VariableID: Unique database identifier
- Variable: VariableName as in ODM
- UnitID: foreign key to units table.
- VariableType: As specified in ODM.
- SampleMedium: As specified in ODM
- IsRegular: As specified in ODM
- TimeSupport: As specified in ODM
- TImeUnitID: foreign key to units table
- GeneralCategory: As specified in ODM
- AltVariableName: VariableName that is being used
- AltVariableCode: VariableCode field as specified in ODM.
- NetworkID: foreign key to HISNetworks table.

MAPPINGSAPPROVED; MAPPINGSNWIS; MAPPINGSCV; MAPPINGSJOINED;

These tables contain concept-variable pairs as defined by data managers and/or the ontology curator.

- ConceptID: The ontology concept which is being mapped.
- VariabeID: foreign key to Variables table.
- DateMapped: Self explanatory.
- DateApproved: <depricated>
- RegisteringIndividual: logged in username of the per
- ApprovingIndividual: <depricated>
- OntologyVersion: Version of the Ontology being used to map.
- MappingID: Unique database identifier

SPATIALREFERENCES AND UNITS TABLES:

Both of these tables and all their fields and values are exactly as they appear in the ODM.

ONTOLOGY TABLES

The CUAHSI Ontology is stored in the database as a set of four primary tables, Concepts, Hierarchy, Synonyms and ConceptPaths. The Concepts table contains the entire list of concepts. Synonyms are concepts with equivalent definitions to terms that exist in the Concepts table. The Hierarchy table maintains the parent/child relationships between the concepts, from which the hierarchy can be derived. The ConceptPaths table is derived from the concepts and hierarchy tables to create a "conceptPath" attribute for each concept. From this attribute, the process of determining the upstream/downstream lineages for each concept is made simple. Upstream is contained in the path, All downstream lineages with contain a similar portion of the path. The field definitions are as follows:

Concepts:

- ConceptID: The ontology concept unique identifier.
- ConceptName: The ontology keyword
- Layer: Number representing the branch level in the ontology tree where the concept resides. 9999 is reserved for leaf nodes.
- IsSearchable: Boolean flag to determine if the concept should be included in the searchable terms list.

Synonyms:

- Synonym: The ontology keyword.
- ConceptID: The ontology concept unique identifier to which this concept is a synonym of.

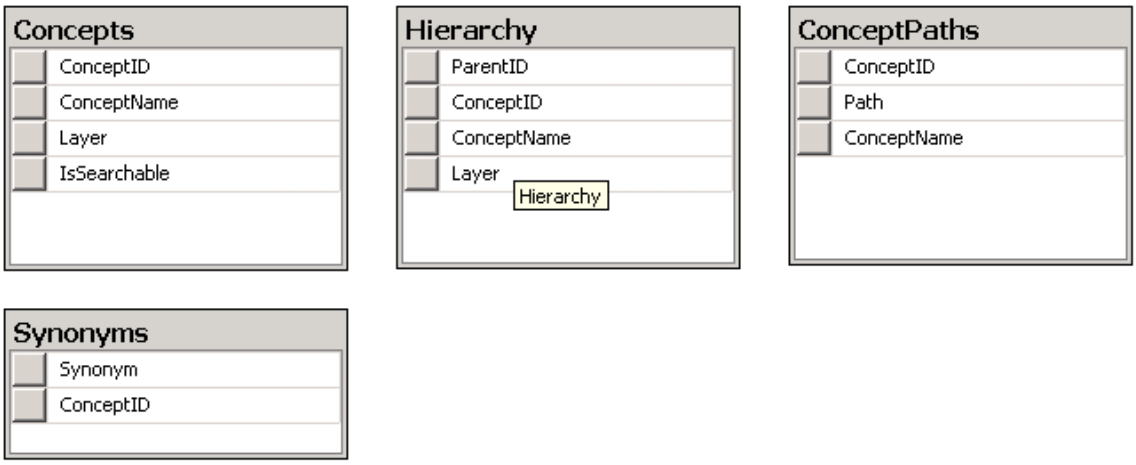
Hierarchy:

- ParentID: The ID of the parent concept. All concepts have a parent except for the "Hydrosphere" root concept.
- ConceptID: The ontology concept unique identifier.
- ConceptName: The ontology keyword
- Layer: Number representing the branch level in the ontology tree where the concept resides. 9999 is reserved for leaf nodes.

ConceptPaths:

- ConceptID: The ontology concept unique identifier.
- Path: The derived attribute illustrating the logical hierarchy.
- ConceptName: The ontology keyword

The figures below picture the four ontology tables, and examples of their content.



Concepts

	ConceptID	ConceptName	Layer
1	1	Hydrosphere	0
2	2	Physical	1
3	3	Level	2
4	4	Level, ice	3
5	5	Ice thickness	9999
6	6	Level, snow	3
7	9	Snow depth	9999
8	10	Snow water equivalent	9999
9	11	Level, stream	3
10	12	Gage height, stream	9999
11	13	Water depth, stream	9999
12	14	Water depth, cross-sectional averaged	9999
13	15	Level, lake	3

Synonyms

	Synonym	ConceptID
1	Depth, snow cover	9
2	Volumetric water content	10
3	Snow water content	10
4	Water content of snow	10
5	Stream gage height	12
6	Stage stream	12
7	Gage height stream	12

Hierarchy

	ParentID	ConceptID
1	1283	1722
2	1283	1723
3	1283	1724
4	1283	1725
5	1283	1726
6	1283	1727
7	1283	1728
8	1283	1729
9	1283	1730
10	1283	1731
11	1283	1732
12	1283	1733
13	1283	1734
14	1283	1735
15	1283	1736
16	1283	1737
17	1283	1738
18	1283	1739
19	1283	1740

ConceptPaths

ConceptID	Path	ConceptName
1		Hydrosphere
2	Hydrosphere/	Physical
3	Hydrosphere/Physical/	Level
4	Hydrosphere/Physical/Level/	Level, ice
6	Hydrosphere/Physical/Level/	Level, snow
11	Hydrosphere/Physical/Level/	Level, stream
15	Hydrosphere/Physical/Level/	Level, lake
19	Hydrosphere/Physical/Level/	Groundwater l
20	Hydrosphere/Physical/Level/	Level, ocean
5022	Hydrosphere/Physical/Level/	Unspecified
24	Hydrosphere/Physical/Area/	Area, ice
26	Hydrosphere/Physical/Area/	Area, atmosph
30	Hydrosphere/Physical/Velocity/	Velocity, strear
32	Hydrosphere/Physical/Velocity/	Velocity, groun
33	Hydrosphere/Physical/Velocity/	Velocity, wind
38	Hydrosphere/Physical/Density/	Water, specific
39	Hydrosphere/Physical/Density/	Density of wat

VIEWS AND INDEXING

The tables described above provide basic information blocks of the HISCentral metadata catalog. In addition HISCentral contains a number of additional views (some materialized) that are used to optimize catalog performance for common service requests, and for commonly requested reports (e.g. on network statistics). Other views are used for maintenance purposes, and support several models of catalog updates as described in a separate section further in the document. While the views listed here don't belong to the set of the core catalog tables, we provide view descriptions here to help an interested reader trace the path from core information blocks to catalog web services, and the implemented optimizations.

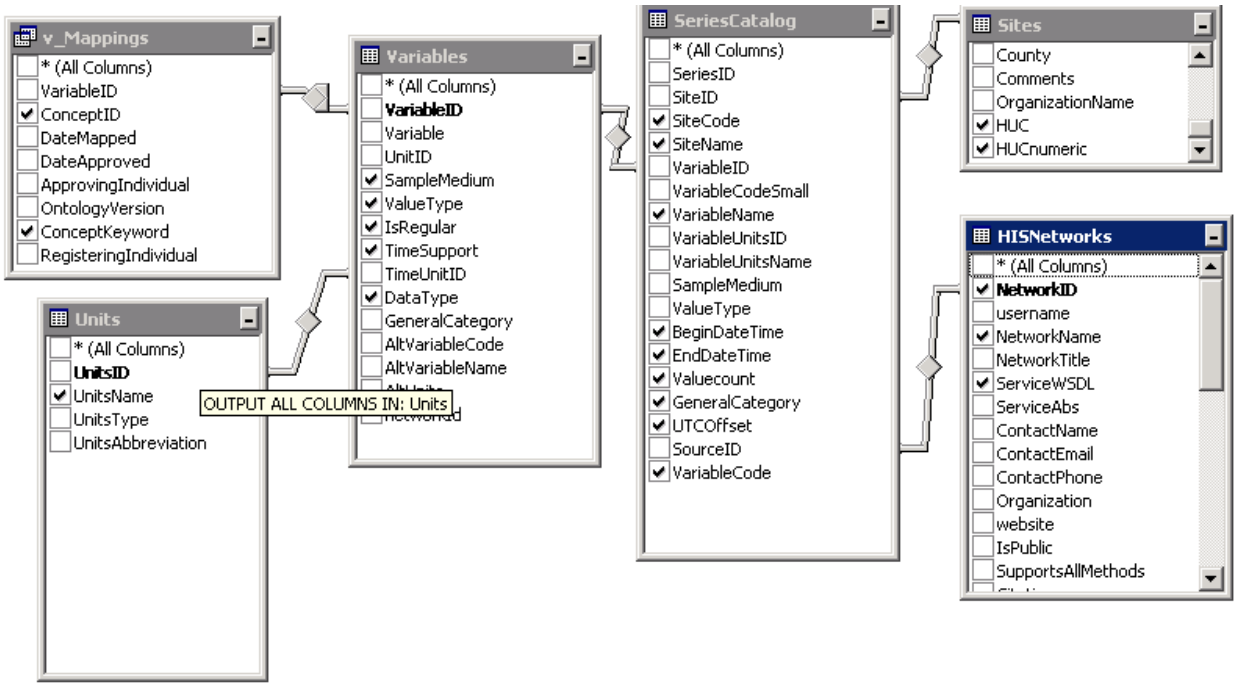
VIEWS USED DIRECTLY BY WEB SERVICES

v_searchableconcepts: Provides a table of terms that can be successfully used to search the catalog. Includes all the synonyms and all concepts that have the "isSearchable" flag set to true.

```
SELECT    ConceptID, ConceptName
FROM      dbo.Concepts
WHERE     IsSearchable = 'True'
UNION
SELECT    ConceptID, Synonym AS ConceptName
FROM      dbo.Synonyms
```

v_conceptsearch: This view is essentially the "flat table view" of the hiscentral catalog, and is designed to be queried on specific columns: Concept keyword, Latitude, Longitude,NetworkID. The resulting table is used to create the SeriesCatalog/Datacart information required by client applications.

Visual diagram of v_conceptsearch view.




```
SELECT     dbo.SeriesCatalog.SiteCode, dbo.SeriesCatalog.SiteName, dbo.Sites.HUCnumeric,
dbo.HISNetworks.ServiceWSDL, dbo.HISNetworks.NetworkName,
                dbo.Sites.Latitude, dbo.Sites.Longitude, dbo.SeriesCatalog.BeginDateTime,
dbo.SeriesCatalog.EndDateTime, dbo.SeriesCatalog.Valuecount,
                dbo.SeriesCatalog.VariableName, dbo.SeriesCatalog.VariableCode, dbo.HISNetworks.NetworkID,
dbo.Units.UnitsName AS TimeUnits, dbo.Sites.HUC,
                dbo.SeriesCatalog.GeneralCategory, dbo.SeriesCatalog.UTCOffset, dbo.Variables.DataType,
dbo.Variables.SampleMedium, dbo.Variables.IsRegular,
                dbo.v_Mappings.ConceptKeyword, dbo.v_Mappings.ConceptID, dbo.Variables.ValueType,
dbo.Variables.TimeSupport
FROM       dbo.HISNetworks INNER JOIN
```

```
        dbo.SeriesCatalog INNER JOIN
        dbo.Sites ON dbo.SeriesCatalog.SiteID = dbo.Sites.SiteID INNER JOIN
        dbo.Variables ON dbo.SeriesCatalog.VariableID = dbo.Variables.VariableID ON
dbo.HISNetworks.NetworkID = dbo.SeriesCatalog.SourceID INNER JOIN
        dbo.Units ON dbo.Variables.TimeUnitID = dbo.Units.UnitsID LEFT OUTER JOIN
```

site_concept_search: Based of the same view for v_conceptsearch, this view is essentially a sub-set of that view. The reason being that to go from concept keyword, to site requires linking the concept to variable to series to site to obtain the relavant data.

```
SELECT    SiteCode, SiteName, HUCnumeric, ServiceWSDL, NetworkName, Latitude, Longitude, NetworkID, ConceptID
FROM      dbo.v_conceptsearch
```

V_Mappings: This view replaces the function of the MappingsApproved table in previous version. It does this by unioning the MappingsApproved with two new tables that contain mappings that were done (to account for the fact that mappings are obtained from several sources: they can be defined by data managers publishing their services, or imported from mapping tables developed for certain large data stores (e.g. USGS NWIS)

```
SELECT    VariableID, ConceptID, DateMapped, DateApproved, ApprovingIndividual, OntologyVersion, ConceptKeyword,
RegisteringIndividual
FROM      dbo.MappingsApproved
UNION
SELECT    VariableID, ConceptID, DateMapped, DateApproved, ApprovingIndividual, OntologyVersion, ConceptKeyword,
RegisteringIndividual
FROM      dbo.MappingsJoined
UNION
SELECT    VariableID, ConceptID, DateMapped, DateApproved, ApprovingIndividual, OntologyVersion, ConceptKeyword,
RegisteringIndividual
FROM      dbo.MappingsNWIS
```

VIEWS USED FOR MAINTENANCE PURPOSES:

The following views are used for ontology-mapping updates:

- **V_mappingsCV:** Is used to create the MappingsCV table. Joins the current version of the Controlled Vocabulary-Ontology mappings to the variables table using the variable name.
- **V_mappingNWIS:** Is used to create the MappingsNWIS table.
- **V_NWIScodes:** is used to correctly handle NWIS codes of different length
- **V_MappingsJoined:** does inner join of variable and concept tables on ConceptName and AltVariableName, to map those variables that match concept names verbatim.

The following views are used to calculate network statistics:

- **V_stats_series**: Tabulates the number of series records and sums the total value count for a network.
- **V_stats_sites**: Tabulates the number of sites and the geographic envelope using the latitude and longitude values for a network.
- **V_stats_variables**: Tablulates the number of variables in the network

- **V_stats_all**: Joins all the other stats views. Used to update statistical fields of the HISNetworks table.

VIEWS USED BY HYDROSEEK:

- **Sources**: A view of essentially the HISNetworks table, but it is filtered using the IsPublic flag in the HISnetworks table to expose only public service on the map.
- **SearchView**: Used in the support HydroSeek search queries.
- **SiteDetailsView**: Used to generate the site details (available variable list) when selecting a site in HydroSeek.
- **MySelect**: Used by HydroSeek tools to manage session selection states.

CATALOG UPDATES FOR FEDERAL DATA SOURCES

Updating the sections of the catalog that deal with federal data sources happens at varying intervals, and depends on whether the catalogs can be harvested directly using web services, or requires a database dump of agency data (or catalog).

Typically, the federal repository catalog harvester is a custom program which executes harvesting from HISCentral using ODM views specialized for that federal source, and will write updates directly to the hiscentral catalog. The data flow is similar to that of the standard harvester described in the next section, but instead of using the web services that connect to a remote HydroServer or an HISServer, it will use direct database connections to a local database that contains a replica of the agency database. This program will be run when the source data is updated. The data workflow is therefore:

- The program queries the ODM views to obtain a list of all sites in the data set.
- For each site, it is compared to what is in the catalog.
- If the site exists already, the catalog site table is updated, else a new record is inserted.
- Again for each site, a list of all time series data sets associated with the site is obtained. From this information, the seriesCatalog and Variables tables are also updated or possibly inserted new.

The above description reflects the workflow within the HISCentral catalog update application. Often, the bulk of work is outside HISCentral's boundaries, and is centered on creating and processing a replica database that reflects an agency catalog (whether it is obtained as a database dump, or as a result of metadata harvesting), and mapping the database fields to ODM fields. This often includes multiple iterations with agencies, as their databases structures and business rules also change. An example of the workflow (for the EPA section of the catalog) is provided below.

PROCESSING THE EPA CATALOG

1) A database dump is provided by EPA. It contains Oracle tables (64 tables). Note that a more recent EPA dump contained only 50 tables, and was differently structured, which necessitated rewriting of the update procedures
2) These tables are loaded in a local Oracle instance running on a Windows 2003 server (kyle.ucsd.edu) with large amount of disk space (several TB).Large storage capacity is required for migrating Oracle tables to SQL server. Loading the content of the Oracle database (tables only) takes one day.
3) Migration to SQL Server, using standard SQL Server migration utility. The large STORET table, called FA_REGULAR_RESULT (with 78 mil rows and 150 columns) takes about 5 days to load (if the load doesn't fail in the process). The content of the table is listed in Appendix B. All Oracle tables are migrated as they are, so that to adjust them or develop views later in SQL Server.
4) Mapping these tables to HISCentral catalog tables. The essential tables to perform these mapping are DI_XXXXXX tables (dimension tables) and FA_XXXXXX tables (fact tables). Of the fact tables, we use the FA_REGULAR_RESULT and FA_STATIONS. Additional metadata required by the catalog, is contained in the dimension tables (activities, intent, medium, matrix, other characteristics); there are 15 DI tables.
5) All the tables mentioned above are used to: a) create a table of sites (from the fact table of stations), b) extract information about source, organization, and other characteristics expected by ODM, from respective fact and dimension tables, c) compute series start and end dates (from the main fact table), and create series catalog. The latter takes one day, once all views and temporary tables are defined and created.
6) Currently, 20 views are used to do the mappings. The key views, for generating catalog's Sites, Variables and SeriesCatalog tables, are shown in Appendix B. Unfortunately, since the structure of the Oracle database changed, the views have to be redone each time.
7) Once this STORET catalog database is created, the HISCentral is updated with the new metadata using the federal repository catalog harvester mentioned above.

CATALOG UPDATES FOR WEB SERVICE DATA SOURCES

The web service harvester is a custom program which connects to the CUAHSI HIS services for each registered data set, and writes updates to the HISCentral catalog. The data flow is the same as in the federal repository catalog harvester, but will connect to the source data through the web services interface. The data is updated in the following workflow:

- The harvester program is passed a list of networkIDs which require updates. This happens on schedule, once a week.
- For each network, the harvester calls the getSites() method on the service to obtain a list of all sites in the data set.
- For each site, it is searched in the catalog so see if it exists.
- If the site exists already, the catalog site table is updated with the new information; else a new record is inserted.
- For each site, the getSiteInfo() method is called to obtain a list of all time series data sets associated with the site is obtained. From this information, the seriesCatalog and Variables tables are also updated or possibly inserted new.

ONTOLOGY UPDATES

The ontology tables are updated manually. The process essentially removes the old ontology tables and replaces them with new ones. The tables updated are: Concepts; Hierarchy; and Synonyms, and the ConceptsPaths table is regenerated based on the three tables. Prior to replacing the ontology tables, changes are verified to learn what existing variable-concept mappings might be affected. To do this, the new concepts table is joined to the existing concepts table and the **v_Mappings** view. This resulting tabular view is then manually inspected, making note of the changes between the two versions of the ontology, and seeing which changes might alter existing mappings. Once it is verified that the updated ontology will function correctly, the tables are renamed so that the new ontology is used. The last step is to recreate the ConceptsPaths table. To do this a custom program is run that calculates the tree path for each concept. The derived path is used to facilitate queries which navigate the tree hierarchy.

CONCEPT MAPPING UPDATES

An ongoing effort in the hiscentral catalog is increasing the volume of data that is directly discoverable through ontology keyword searches. In addition to variables being manually mapped to concepts using the HydroTagger application in HIS Central, variables are now also being mapped administratively (specifically, this process is adopted for federal data sources). All HydroTagger mappings are stored and managed in the MappingsApproved table of the hiscentral catalog. The additional mapping tables integrated with the MappingsApproved table are: MappingsNWIS, MappingsCV, and MappingsJoined: To provide access to all the mappings, a database view has been created, **v_Mappings,** which uses SQL union statements to combine the multiple mapping tables.

- MappingsApproved: This is the original mapping table and is used by the HydroTagger application to map variables.
- MappingsNWIS: This table is the result of an ongoing effort to map all of the NWIS 5 digit numeric codes to ontology concepts.
- MappingsCV: Much of the CUAHSI controlled vocabulary has also been mapped, so variables which comply with the CUAHSI controlled vocabulary are possibly mapped.
- MappingsJoined: Variables names which are exact string matches to ontology terms are mapped to that term in this table.

When a new table of mappings is received from the CUAHSI office, the following procedure is followed:

- The NWISParameters_Lookuptable spreadsheet (with three columns: NWISParamCode, NwisParamdesc, and conceptid) is imported in SQL Server.
- A fourth "code" is added to properly format values in the NWIS ParameterCode column (remove spaces and validate)
- A view "**v_NWIScodes**" is created, which selects from the variables table all NWIS variables, and formats the VarCode string to strip off all text other than the 5 digit NWIS code.
- An additional view is created, **v_MappingsNWIS**, which joins the NWISParameters_Lookuptable to v_NWIScodes.

The current status of the mappings is as follows:

- The initial NWISParameters_Lookuptable had 8400 records, roughly 4000 of which were mapped to a valid conceptid (conceptid field is NULL when unmapped).
- The HISCentral catalog contains 980 variables for the 4 NWIS data services combined. Of these 980, 340 were mapped using this table. Only 540 of the 980 NWIS variables are in the NWISParameters_Lookuptable
- Currently, 40.43% of the variables, 63.04% of the series, and 92% of the values referenced in the catalog, are discoverable. A table detailing the current status of the tagging, is in Appendix C.

FUNCTIONAL OVERVIEW OF THE HISCENTRAL APPLICATION

SERVICE REGISTRATION OUTLINE

The HIS Central web interface is designed primarily for data managers who want to register their services and manage information associated with them.

Below we present an outline of a step-by-step process that data managers are expected to follow when registering their data in the HIS Central. Some steps are self-explanatory, and follow common interface practice and conventions (such as creating a login account, or requesting forgotten passwords). Other steps that are specific to HISCentral are discussed in more detail in separate sections.

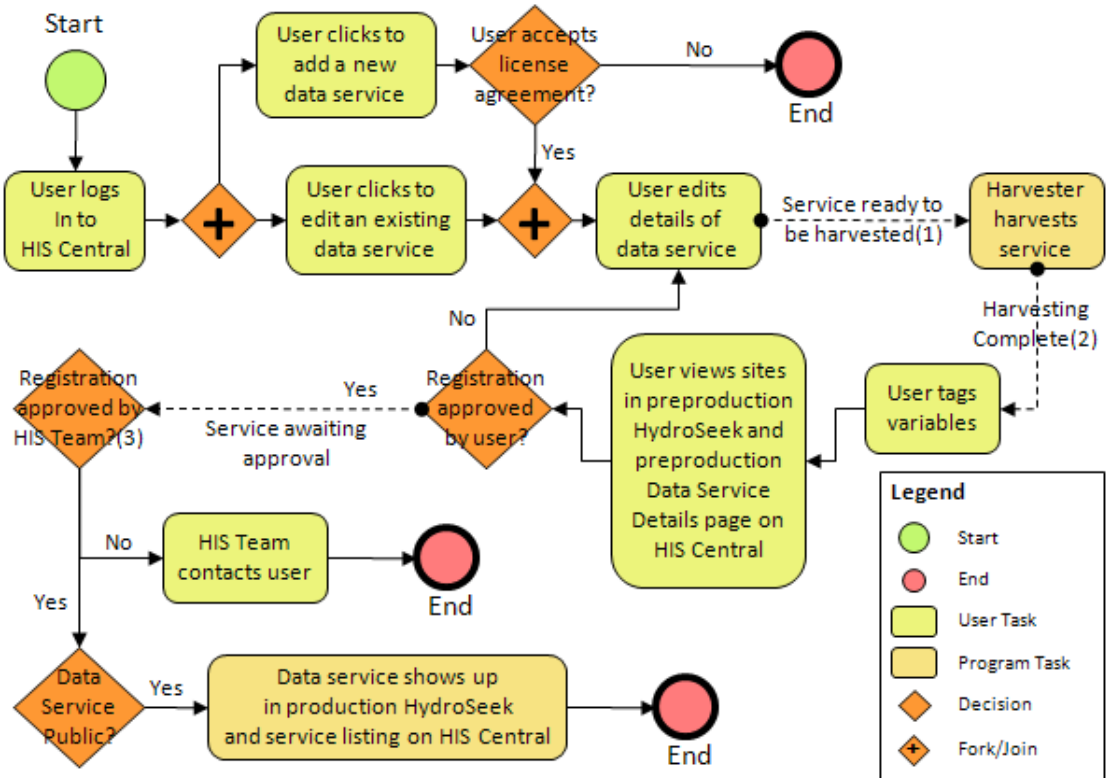
1) Create a login account at http://hiscentral.cuahsi.org
2) Add a new Data Service
3) Test the data service using the test page to make sure it works as you expect
4) Change the status of your service to public.
5) The HIS Central Administrator will trigger a harvest of your service.
6) You will receive an email notification of the results of the harvest and will be prompted to return to the HIS Central site to tag the harvested variables.
7) Once your variables are tagged, you can test the concept search functionality in the HydroSeek application, using a pre-publication (test) version (http://test.HydroSeek.net/search)
8) The updated metadata catalog will be pushed to the production version of HydroSeek (www.HydroSeek.net) at the main HIScentral site at SDSC, and replicated off site at the University of Texas.

PREPRODUCTION AND PRODUCTION SUBSYSTEMS

Quality control should be performed before a newly registered or edited data service shows up in HydroSeek. To facilitate quality control, each data service first shows in a **preproduction** part of the system, before it is displayed in the **production** part. Once a service is registered and variables are tagged, the user who registered the service will open the preproduction HydroSeek to check that sites and variable listings are showing up in a satisfactory

manner. If so, the user submits the service to the HISCentral administrator for adding to the production system. Once a service is approved, it is then classified as "production", and if public, will then be viewable by regular users of HIS Central and also in the production HydroSeek.

A diagram outlining service publication, testing and approval steps is shown in the figure below, and each step in the process is detailed in the following sections.



NOTES:
(1) For 1.1, notification is sent to administrator, who then triggers harvest manually.
(2) All notifications to user are sent to the email address the user provided upon registration.
(3) "HIS Team" refers to a member of the HIS Team assigned to evaluate the data service for approval.

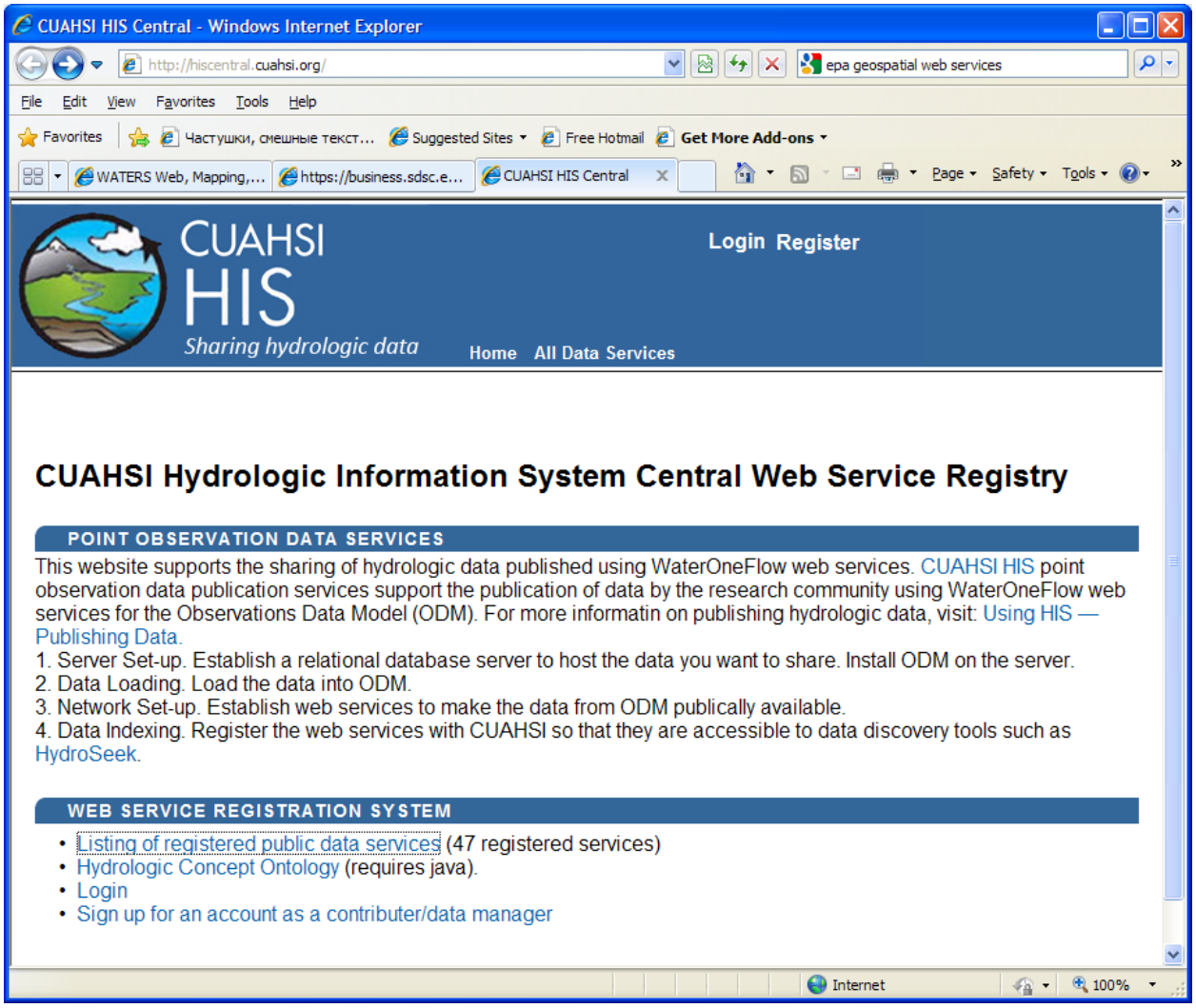
Functionality of HISCentral in both open and authenticated modes is described below.

HISCentral functionality not requiring user authentication

In this mode, HISCentral visitors can browse the list of publicly registered services, and view metadata and statistics for each service.

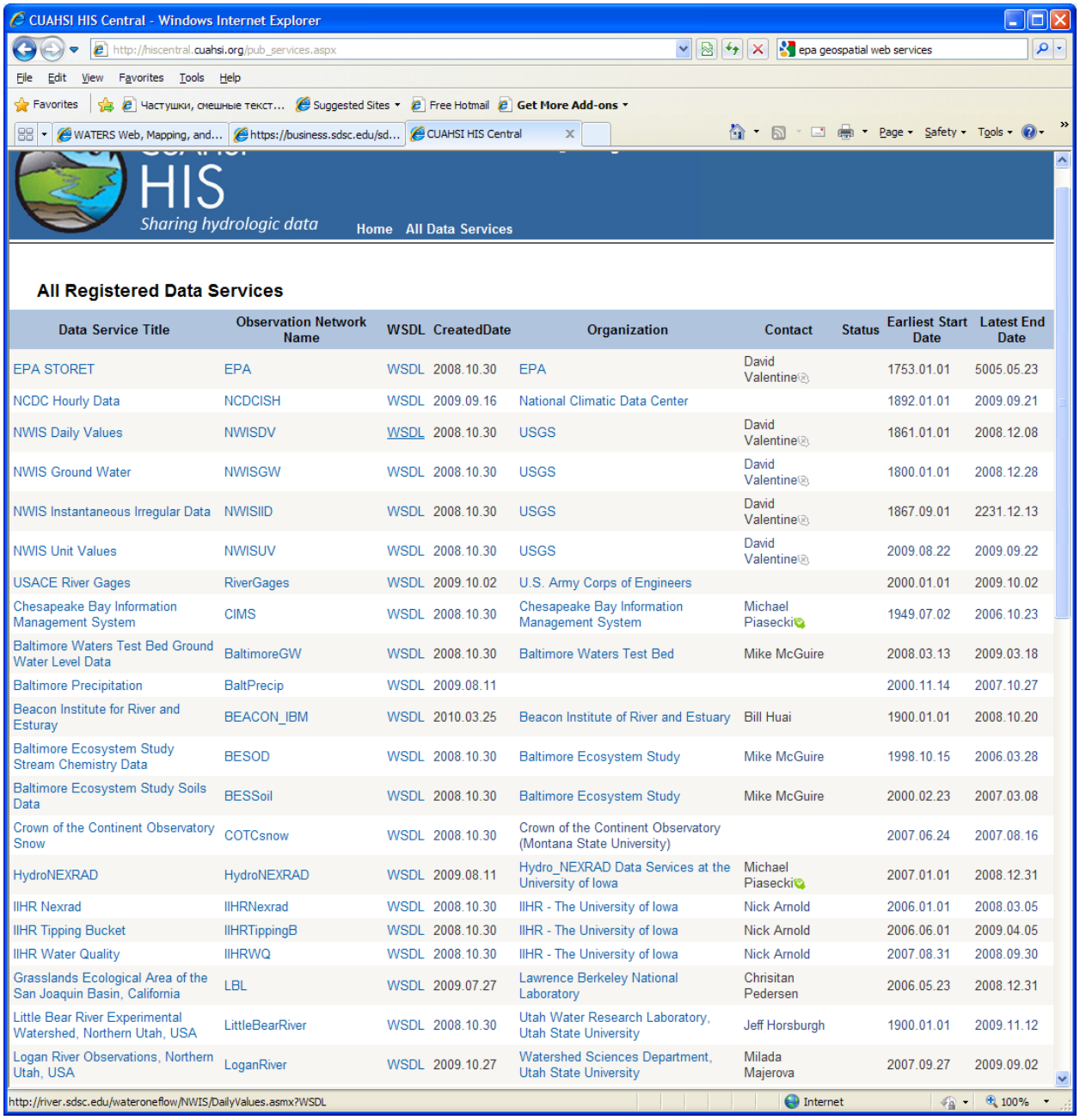
Main Page:

This page provides the main entry points into the system, points the user to the main applications of the HISCentral, the ontology browser, and the help system.



PUBLIC SERVICES LIST PAGE:

Each registered service marked as public will be displayed in a tabular HTML page. For each entry, the service's Name, Title, WSDL and contact information is displayed. The service name column will link users to the service details page for that service. Services from federal agencies are shown at the beginning of the list.



SERVICE DETAILS PAGE:

The service details page displays all the information provided for each registered service. Additionally, statistics about the service are also displayed. The rubrics on the page include: service name, web address and contact person information; service statistics (number of sites, variables, values; spatial extent); Abstract; Citation (a reference to an appropriate publication, or an acknowledgement recommended by the data publishers, which data users are encourage to reference when publishing any work where the contributed data were used); Geographic description; Keywords; Study description; Spatial coverage; Temporal coverage.

CUAHSI HIS Central - Windows Internet Explorer

http://hiscentral.cuahsi.org/pub_network.aspx?n=52

File Edit View Favorites Tools Help

Favorites | Частушки, смешные текст... | Suggested Sites | Free Hotmail | Get More Add-ons

WATERS Web, Mapping, and... | https://business.sdsc.edu/sd... | CUAHSI HIS Central

Find: hors Previous Next Options

CUAHSI HIS

Sharing hydrologic data

Login Register

Home All Data Services

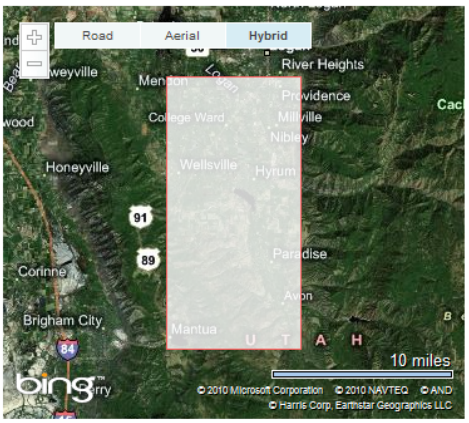
Little Bear River Experimental Watershed, Northern Utah, USA

Utah Water Research Laboratory, Utah State University

LittleBearRiver

http://icewater.usu.edu/littlebearriver/cuahsi_1_0.asmx?WSDL

Contact: Jeff Horsburgh
jeff.horsburgh@usu.edu
435-797-2946



Service Statistics:

Sites: 14
Variables: 58
Values: 6000695

Geographic Extent:

41.71847
-111.9464 -111.7993
41.49541

Abstract

Utah State University is conducting continuous monitoring within the Little Bear River watershed of northern Utah, USA to investiage the use of surrogate measures such as turbidity in creating high frequency load estimates for constituents that cannot be measured continuously.

Geographic Description:

The Little Bear River in northern Utah, United States is typical of many semi-arid watersheds in the western United States where stream discharge is dominated by spring snowmelt runoff and where extensive hydrologic modification for agricultural diversion has taken place. The Little Bear River is a part of the larger Bear River system and terminates at Cutler Reservoir, a shallow, eutrophic reservoir built on the mainstem of the Bear River, which ultimately drains to the Great Salt Lake. The Little Bear River watershed encompasses approximately 740 km2 of primarily agricultural and forest lands. Approximately 70% is grazing land and forest, 19% is irrigated cropland and 7% is dry cropland. The area is experiencing rapid population growth, with a 32% increase in population between 1990 and 2000. Elevation in the watershed ranges from 1,340 m to over 2,700 m and most of the annual precipitation falls as snow at higher elevations, with occasional summer storms. A reservoir traps high mountain runoff in the East Fork, and another reservoir is located below the confluence of the branches of the river. In both cases, these reservoirs release water during the summer irrigation season, which generally spans from mid April to late September, to satisfy agricultural diversions that essentially dewater portions of the river below the reservoirs during much of the irrigation season.

Citation

Horsburgh, J. S., D. K. Stevens, D. G. Tarboton, N. O. Mesner, A. Spackman Jones, and S. Gurrero (2009) Monitoring data collected within the Little Bear River Experimental Watershed, Utah, USA, Utah State University.

Done | Internet | 100%

HISCentral functionality that requires user authentication

In this mode, users can create and manage the entries for (i.e., register) water data services, upon authenticating to HISCentral. The description here follows the general outline of web service registration steps as presented in the "Service Registration Outline" section

1. User authentication:

At this step, user enters login information to get access to service management functionality of HISCentral. Users can request account from the main HISCentral page. Password request service is also supported.

Accepting the Data Publishing Agreement

Once they have logged in, users will be able to register new data services and edit the description of existing data services that they have created through forms within HIS Central. As part of the registration process, users are prompted to accept the data publishing agreement, as linked on the HIS website at http://his.cuahsi.org/governance.html. Users must accept the agreement in order to continue the registration process.

2. Adding a new service, and editing service details:

Adding a new data service to the metadata catalog requires that the user fills out a standard web based form. When adding a new service, the initial page only requires a subset of the fields:

- Service Name: this should correspond with the "Network" name, the value the data manager provided when configuring a WaterOneFlow web service against an ODM data base. Service Title
- Service WSDL. The web access point to your data server. Be sure to provide a URL that is accessible from the internet.

Once the submit button is clicked and the service is added, the page is directed to the service details editing page (see the figure below in this section). On this page, users can add additional service information into a form that contains the following fields:

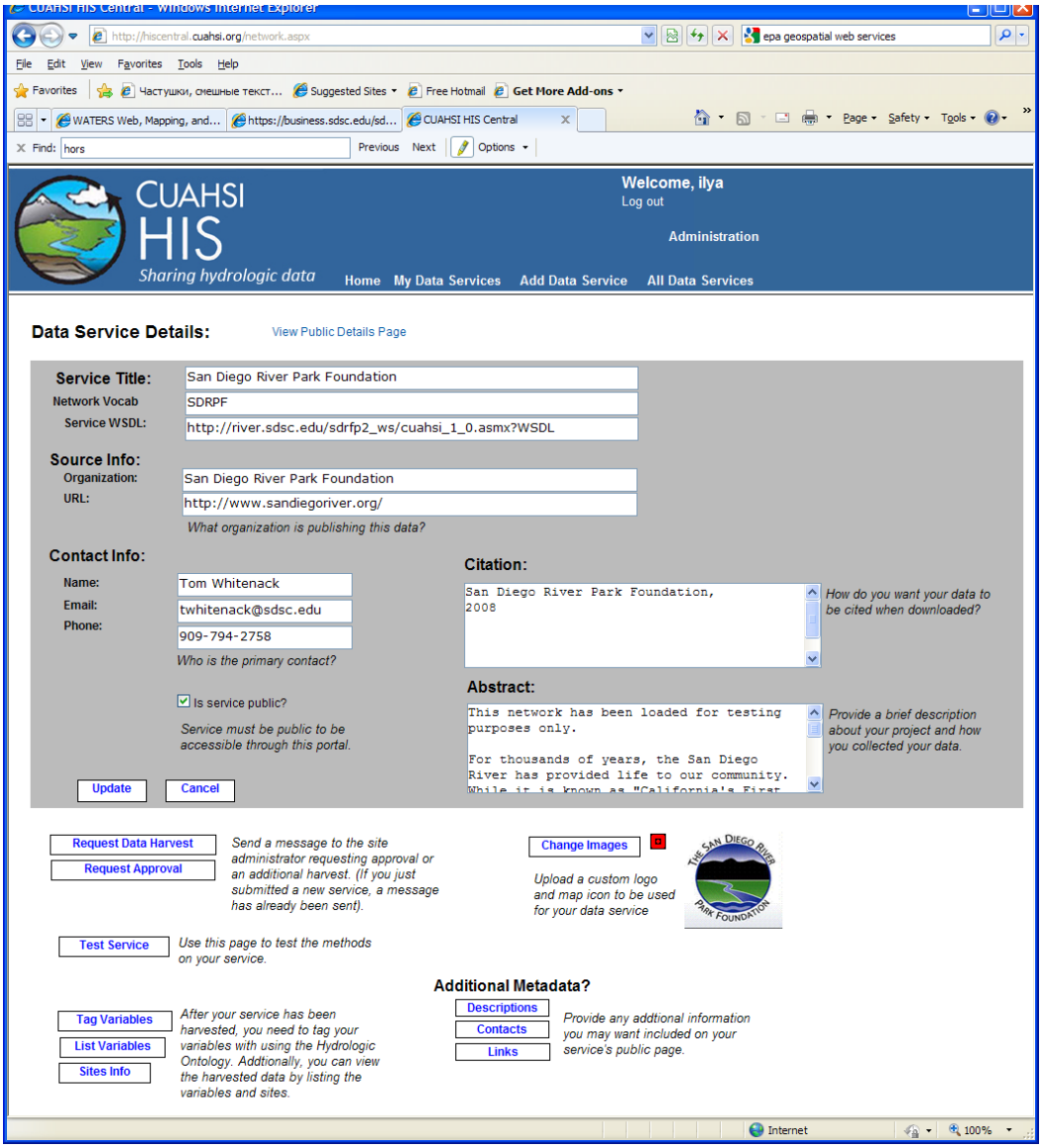
- Source Info: What is the name of the organization that is responsible for the data service? Do they have a web site?
- Contact Info: Who should be contacted with questions regarding this service? How? This contact is considered the primary contact.
- Citation: What should the citation text read in the files downloaded from your service?
- Abstract: A general description of service data. Where it came from, why was the study done, etc.
- Is the service public? Unchecked, or private by default, this provides you with the ability to take your service offline from our system. Services which are tagged private, aren't visible to other users and aren't included in HydroSeek queries.

Custom Image Uploads:

- Icon: This is the image that will be displayed on a map to represent the sites from the observation network. This image should be roughly 15 x 15 pixels
- Logo: This is primarily for branding purposes. In the HydroSeek client application, when a user clicks on a site, the dialog that appears will have the organization logo on it. This image should be roughly 150 x 150 pixels.

Additional Metadata: In case the existing fields aren't adequate to describe and document a service, an expandable set of additional fields is available, in the following rubrics:

- Contacts: Any additional personnel can be added.
- Links: Any additional web resources, whether they be online documents or external web sites can be included.
- Descriptions: Any additional descriptive information can be included here.



Service metadata fields are summarized in the following table:

Attribute	Description	Example
Network Vocabulary	The name of the vocabulary used by the data service. Used to call the web services, i.e., as the Vocabulary prefix in "Vocabulary:VariableCode".	LBR
Data Service Title	A descriptive title for the data service.	Little Bear River Test Bed, Utah, USA
Data Service WSDL Address	The URI for the data service WSDL.	http://his02.usu.edu/littlebearriver/cuahsi_1_0.asmx?WSDL
Data Service/Project URI	URI to a site that gives additional information about the data service. This may be a different URI than those listed in the SourceLink element of WaterML (equivalent to the SourceLink attribute in the Sources table of ODM).	http://water.usu.edu/littlebearriver/
Project Status	An indication of whether the published data are historical or if data collection is ongoing. To be chosen from a list of potential status.	Historical Ongoing Realtime
Responsible Organization	The organization that is responsible for maintaining the data service. This may not be the same as the Organization child element of elements of type SourceType in WaterML (equivalent to the Organization attribute in the Sources table of ODM).	Utah Water Research Laboratory, Utah State University, Logan, UT, USA
Organization URI	URI to a site that gives additional information about the organization that is responsible for maintaining the data service. This may be a different URI than those listed in the SourceLink element of WaterML (equivalent to the SourceLink attribute in the Sources table of ODM).	http://uwrl.usu.edu
Technical Contact Information	Contact information for the manager of the data service. Likely not the same as the contact information listed in elements of type SourceType in WaterML (equivalent to information in the Sources table in ODM). This is a data service level contact and not a data source level contact.	Name: Jeff Horsburgh Address: 8200 Old Main Hill, Logan, UT 84322-8200, USA Phone: (435) 797-2946 Email: jeff.horsburgh@usu.edu NOTE: Implemented as separate fields for name, address, phone, email, etc.
Lead Scientist	Contact information for the lead scientist or project principal investigator in charge of interpreting or using the data. May not be the	Name: Jeff Horsburgh Address: 8200 Old Main Hill, Logan, UT 84322-8200, USA

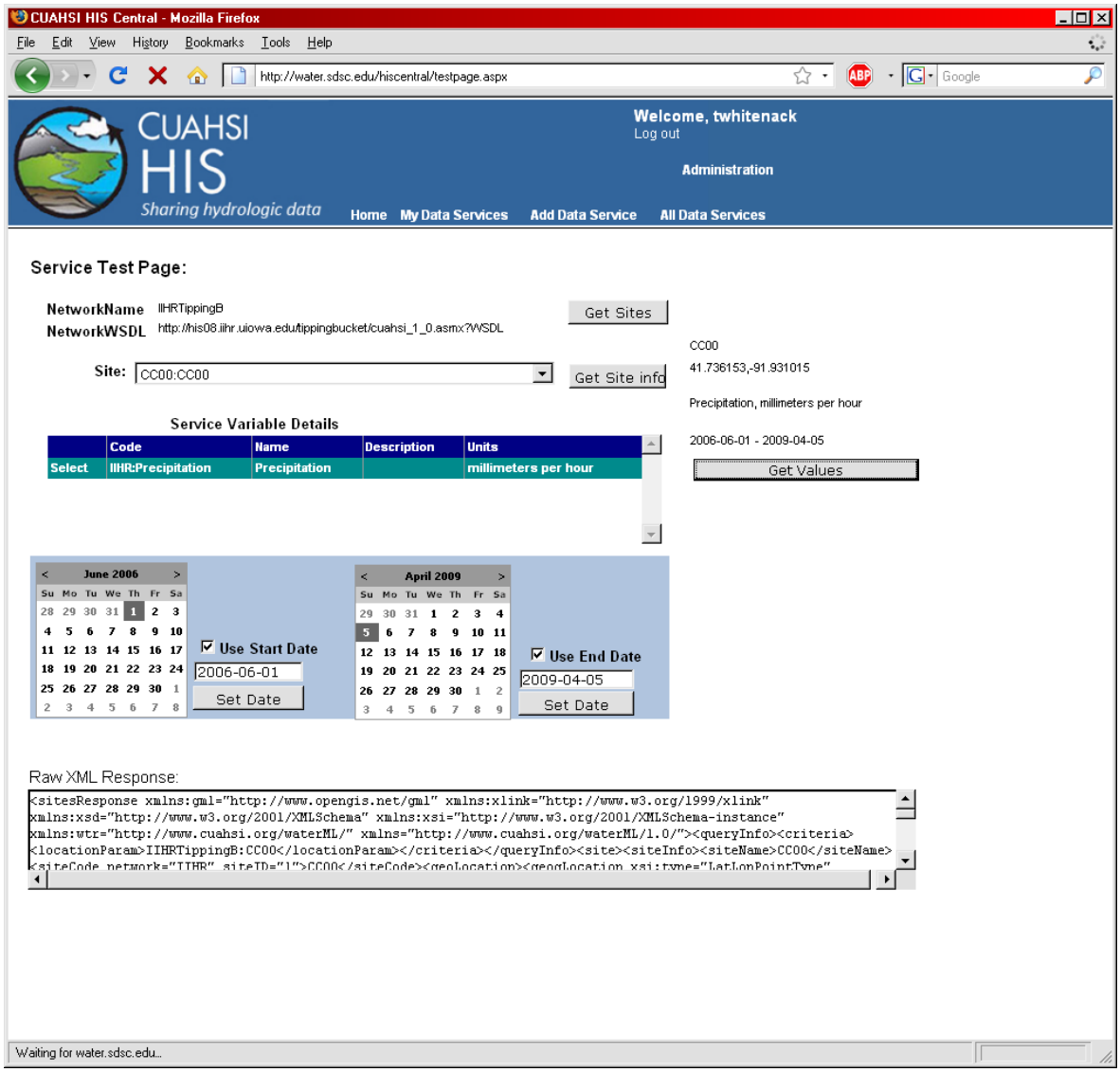
	same as the information listed in elements of type SourceType in WaterML (equivalent to information in the Sources table in ODM) because there may be multiple sources represented by data in the data service. This is a data service level contact and not a data source level contact.	Phone: (435) 797-2946 Email: jeff.horsburgh@usu.edu NOTE: Implemented as separate fields for name, address, phone, email, etc.
Descriptions	A table that lists description categories and then the description itself. The table permits data publishers to provide information for which a formal HIS Central attribute does not exist. The default description category will be "General." The data publisher may enter as many description categories as desired.	**Category: Description** General: We study short-term var… other examples Study Area: Little Bear River, which drains… Sensor Array: We use 12 sensors along… Physical processes: Dissolved Oxygen dynamics, stream temperature, pollutant loading … Geographic Description: The Little Bear River located in northern Utah… History: Data collection began with streamflow in 2005. Other data was added …
Documents and Publications	A list of publications related to the data. These are input into a table with the following fields: Document Citation Document URI	Horsburgh, Tarboton and Stevens, 2008, A Model for Turbidity in the Little Bear River, Water Resources Research, 45: W08734, doi:10.1029/2007WR006392 http://www.lbr.com/paper1/
Links	A list of additional websites related to the data. These are input into a table with the following fields: Link Description Link URI	Little Bear River Project Homepage http://www.lbr.com
Data Service Logo	The logo that is associated with the data service and that is displayed in applications such as HydroSeek. To be uploaded by the user upon registration.	
Data Service Icon	The icon used as the marker for sites within the data service within mapping applications such as HydroSeek . To be uploaded by the user upon	

	registration.	
Data Service Citation	The data service level citation to be used when referencing or acknowledging use of the data (not to be confused with data source citations in the citation child element of elements of type SourceType in WaterML (equivalent to information in the Sources table from ODM).	Horsburgh, J. S., D. K. Stevens, D. G. Tarboton, N. O. Mesner, (2008), Continuous discharge, water quality, and meteorologic data collected in the Little Bear River, UT, USA, http://his02.usu.edu/littlebearriver/cuahsi_1_0.asmx?WSDL.
Public or Private	An indication of whether or not the data service is to be public. If public, data will be accessible from HydroSeek, and the service will appear in the list of registered public data services on HIS Central.	

3. TESTING A SERVICE

We strongly recommend that users test their service thoroughly upon registering it at the HISCentral. To do this, click the "Test" link, to retrieve sites, variables and data series, and eventually data values, and explore them for correctness (see figure below).

In addition, from the Data Service Details page the user can list Sites and Variables, and explore SeriesCatalogs for each selected site.



4. MAKING THE SERVICE PUBLIC

Once the data publisher tested the service and is comfortable with its performance, she can request that the service be harvested. . Then to make the service public, place a checkmark in the “public” checkbox in the Data Service Details page.

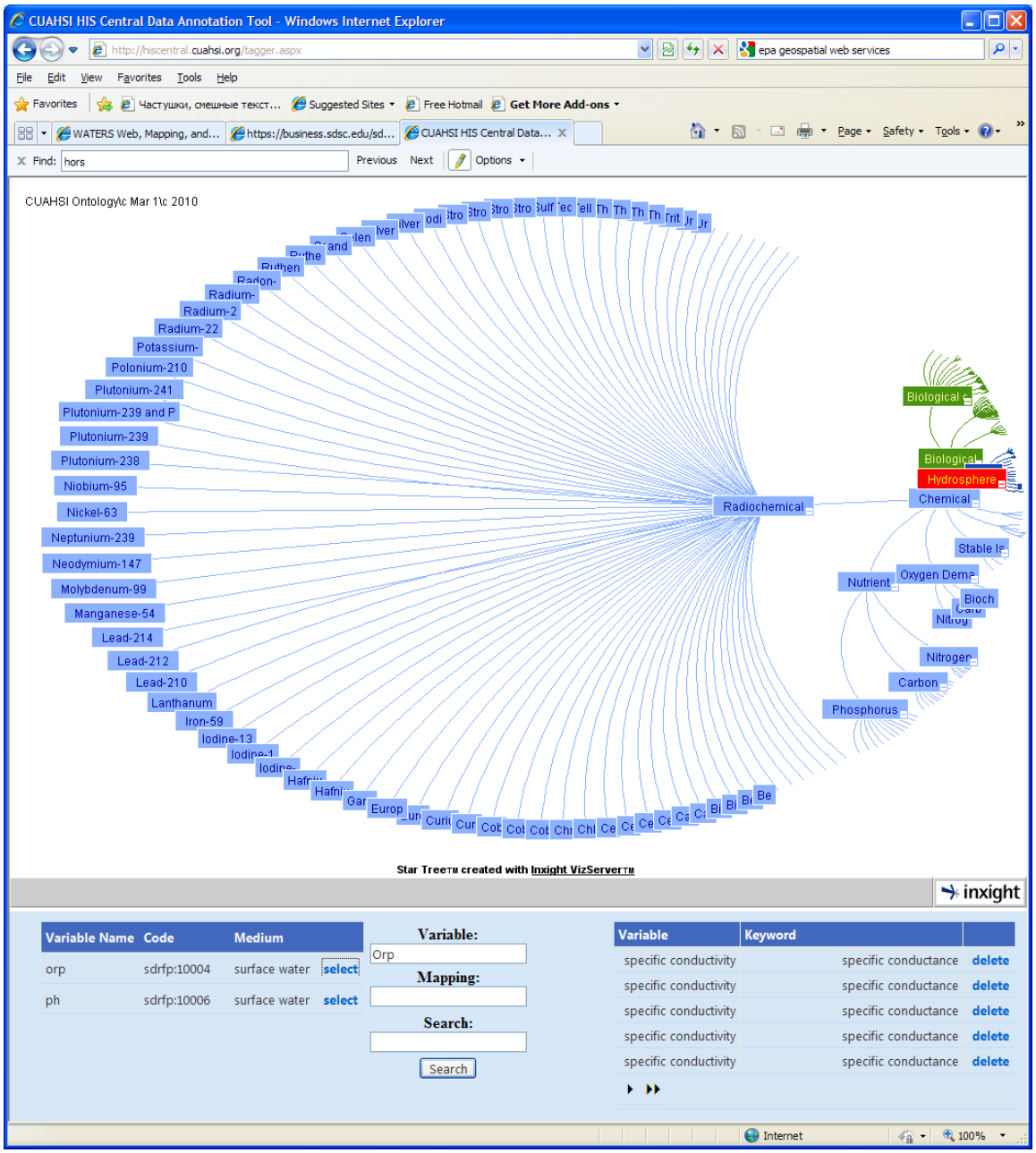
5. SERVICE HARVESTING

Once a WaterOneFlow service is registered as public, its series information is harvested using the service’s GetSites, and GetSiteInfo methods, which are available to all clients of WaterOneFlow web services. It starts with a call to the getSites() method, which predictably returns a list of every site within the service. Then the harvester process iterates through each site, obtaining the variable information as well as details about the data series collected at each site. Harvesting is a CPU intensive process for both the remote server as well as the metadata base server. The harvested metadata are added to the central metadata catalog. Once the initial harvest is completed, the data

service will be re-harvested on a weekly basis. Special care is taken to assure that only new sites and new variables are added. All information about the service (as described in earlier service registration pages), plus summary of harvested data (e.g. spatial and temporal extents of the service, counts of sites and variables, etc.), as well as the date of last metadata harvest, is preserved in the "HISnetworks" and related tables of the central metadata catalog.

6. TAGGING VARIABLES

In order to support the search capabilities by establishing the necessary concept ID and variable ID linkages the semantic mediation framework contains another application named HydroTagger. HydroTagger provides a graphical interface in which one can search for the appropriate concept to tag a variable against. HydroTagger operates off the same database tables as HydroSeek and has the main task of managing tagged and non-tagged variables. Non-tagged variables are "discovered" by a crawler that, currently once a week, trawls through all registered web services marked for harvesting (those of the large nationwide and regional data sources are updated less frequently because they are very work intensive) to find out what has been added in the last week. Before you can start tagging variables, you will receive an email from the HISCentral indicating that the harvesting of your service metadata has been complete, and prompting you to start variable tagging. The tagging process is designed to be intuitive, and is also described in the help files of the application.



HISCentral Web Service

The HISCentral web service exposes the content of the central service registry and metadata catalog. The service is available from http://hiscentral.cuahsi.org/webservices/hiscentral.asmx, and includes the following methods:

Ontology Service Methods

GetSearchableConcepts(): Provides a list of all searchable keyword concepts.

- Required parameters: none.
- Returns: (array of string) a list of all searchable concepts from the HIS ontology. Searchable concepts include “branch” concepts as well as “Leaf” concepts. Higher level branches are not included as they are too broad

GetOntologyTree(): Provides a portion of a the ontology tree, as a tree structure.

- Required parameters:
  - Keyword (string) - must match existing keywords
- Returns: (OntologyNode) a tree of concepts in XML format, starting with the input concept as the root and including its child nodes. Passing in a “Branch” concept. (Passing “HydroSphere” returns then entire ontology).

GetWordList(): Provides a list of keywords which match a character sequence passed to it.

- Required parameters:
  - PrefixText - (string) text to search on May be prefix or a portion of text within the keyword (searching for “flow” will return “streamflow”).
  - Count - (int)max number of results to return
- Returns: (array of string) a list of keywords which contain the passed in search text. This method would provide “intellisense” for keyword entry box, to steer clients towards the known keywords.

HISCENTRAL CATALOG SERVICE METHODS

GetWaterOneFlowServiceInfo(): Provides a list of all the services registered with HIS Central.

- Required parameters: None
- Returns: (array of ServiceInfo) a list of registered services with service metadata as available in the HISNetworks table

GetServicesInBox(): Provides a list of all the services which overlap a specific region.

- Required parameters:
  - Box – geographic bounding box to search in.
- Returns: (array of ServiceInfo)a list of registered services.

GetSitesInBox (): Provides the site information necessary to display sites on a map and request more information about series.

- Required parameters:
  - Bounding Box (box)- The geographic envelope in which to search

- Concept Keyword (string) – This is changed from previous versions which used the "conceptCode" Keyword passed must match an existing searchable keyword, or be an empty string.
- NetworkIDs (array of int)- The HISCentral NetworkIDs that will be included in the search. Passing an empty array will search all registered networks.

- Returns: (array of site) list of all sites that fall within the bounding box, have variables that are mapped to or fall under the Ontology Concept keyword, and are within the list of services)

GetSeriesCatalogForBox (): This is primary method for searching the catalog. It provides series record information, which is then used by the client to create data carts and obtain data values from the registered service.

- Required parameters:
  - Bounding Box (box)- The geographic envelope in which to search
  - Concept Keyword (string) – This is changed from previous versions which used the "conceptCode"
  - NetworkIDs (array of int)- The HISCentral NetworkIDs that will be included in the search. Passing an empty array will search all registered networks. \
  - BeginDate (string) – formatted " 1/1/2000" or " 01/01/2000"
  - EndDate (string)
- Returns: (array of seriesRecord)the series catalog within a specified Lat/Lon box which also match the keyword query)
  :

APPENDIX A: LIST OF METHOD PARAMETERS FOR HISCENTRAL SERVICE

LIST OF OBJECTS USED BY THE SERVICE:

- **ServiceInfo :**
  - Fields:
    - ServURL – (String)WSDL endpoint for WaterOneFlow web services
    - Title – (String)service title
    - ServiceDescriptionURL – (String)
    - Name – (String)contact name
    - Email – (String)contact email
    - Phone – (String)contact phone number
    - Organization – (String)name of organization responsible
    - Value count – (Int)total number of values offered by this service
    - Variablecount - (Int)number of variables offered by this service
    - Sitecount – (int)number of sites in this service
    - ServiceID – (int) unique database identifier for this service
    - NetworkName - (String) Name of network, ie "EPA"
    - minx, miny, maxx, maxy – (double) bounding envelope (lat/long) of the service
- **Box**

- Fields:
  - minx– (double) min longitude value
  - miny– (double) min latitude value
  - maxx– (double) max longitude value
  - maxy – (double) max longitude value

- **SeriesRecord**:
  - Fields
    - ServCode - (string) services unique code – “nwis”
    - ServURL – (string) wsdl address of service
    - Location – (string) site code
    - VarCode- (string) variable code associated with the series
    - Varname –(string) variable name
    - beginDate – (string) start date of series
    - endDate – (string) end date of series (as of last harvest).
    - Authtoken – (string) unimplemented
    - ValueCount – (int) number of values in series
    - Sitename –(string) site name
    - Latitude –(double)
    - Longitude – (double)
- **MappedVariable**
  - Fields:
    - variableName – (string) variable name
    - servCode – (string) service code
    - WSDL – (string) service wsdl
    - conceptKeyword – (string) keyword from ontology which variable is mapped.
- **Site**:
  - Fields:
    - SiteName – (string) name of site
    - SiteCode –(string) sitecode
    - Latitude –(double)
    - Longitude – (double)
    - HUC – (string) 8 digit NHDPlus huccode. (may contain leading zero’s)
    - hucNumeric - (int) numeric representation of the huc code
    - ServCode
    - ServURL
- **OntologyNode**
  - Fields:
    - ConceptCode- (string) going away / possibly replaced with conceptid
    - Keyword-(string)
    - childNodes[] – (array of OntologyNode) this node’s child nodes

APPENDIX B: MATERIALS FOR THE EPA CATALOG LOADING EXAMPLE

FIELDS IN THE FA_REGULAR_RESULT TABLE FROM THE ORACLE DUMP:

PK_ISN	NUMBER	No	1	Primary Key. Generated using sequence object SEQ_REGULAR_RESULT.
ORGANIZATION_ID	VARCHAR2(256 BYTE)	Yes	2	The user-defined code that uniquely identifies an Organization. Organization IDs are registered centrally with STORET so that they are guaranteed to be unique on the National level. They become part of the Oracle Key on every row in the STORET database that belongs to the Organization.
ORGANIZATION_IS_NUMBER	NUMBER(20,0)	Yes	3	A system-generated value used to uniquely identify an occurrence of this table.
STATION_ID	CHAR(15 CHAR)	Yes	4	Station identifier.
STATION_NAME	VARCHAR2(256 BYTE)	Yes	5	Station Name
ACTIVITY_START_DATE_TIME	DATE	Yes	6	For Samples Collected or for Measurements/Observations made - the date that the Field Activity began. For Samples Created from other samples - the date on which a sample is created by compositing, splitting, or subsampling from a parent sample.
ACT_START_TIME_ZONE	VARCHAR2(5 BYTE)	Yes	7	Field activity start time zone
TRIP_ID	VARCHAR2(256 BYTE)	Yes	8	Field Trip Identifier
TRIP_NAME	VARCHAR2(256 BYTE)	Yes	9	Field Trip name
STATION_VISIT_ID	CHAR(3 CHAR)	Yes	10	Station Visit identifier
CHARACTERISTIC_GROUP_TYPE	VARCHAR2(256 CHAR)	Yes	11	
CHARACTERISTIC_NAME	VARCHAR2(256 CHAR)	Yes	12	Name of physical/chemical, biological and habitat characteristics. It is also called Substance Name in Substance Registry System.
RESULT_VALUE	NUMBER	Yes	13	Result value
RESULT_UNIT	VARCHAR2(256 BYTE)	Yes	14	Result Unit code
RESULT_VALUE_TEXT	VARCHAR2(256 BYTE)	Yes	15	Result value text
SAMPLE_FRACTION_TYPE	VARCHAR2(256 BYTE)	Yes	16	Sample Fraction Type
RESULT_VALUE_TYPE	VARCHAR2(256 BYTE)	Yes	17	Result Value type

STATISTIC_TYPE	VARCHAR2(256 BYTE)	Yes	18	A statistic or calculation type which describes the reported result (e.g., average, mode, median, MPN).
RESULT_VALUE_STATUS	VARCHAR2(12 CHAR)	Yes	19	Result value status
WEIGHT_BASIS_TYPE	VARCHAR2(256 BYTE)	Yes	20	Weight Basis type
TEMPERATURE_BASIS_LEVEL	VARCHAR2(256 BYTE)	Yes	21	A statistic or calculation type which describes the reported result (e.g., average, mode, median, MPN).
DURATION_BASIS	VARCHAR2(256 BYTE)	Yes	22	Duration Basis
ANALYTICAL_PROCEDURE_SOURCE	VARCHAR2(256 BYTE)	Yes	23	The name of the Analytical Procedure Owner. Examples include the American Society for Testing Materials, the United States Geological Survey, and the Environmental Protection Agency.
ANALYTICAL_PROCEDURE_ID	VARCHAR2(256 BYTE)	Yes	24	The abbreviated name or identifying code of the analytical procedure.
LAB_ID	CHAR(8 CHAR)	Yes	25	Laboratory Identification
LAB_NAME	VARCHAR2(60 CHAR)	Yes	26	Laboratory Name
LAB_CERTIFIED	VARCHAR2(1 BYTE)	Yes	27	Laboratory Certification information
LAB_BATCH_ID	CHAR(12 CHAR)	Yes	28	The code that represents the laboratory batch ID for a result or a group of results. Batch ID may be used to link with Laboratory Data files, which may contain useful information further describing the reported result.
ANALYSIS_DATE_TIME	DATE	Yes	29	Sample Analysis Date and time stamp
ANALYSIS_TIME_ZONE	VARCHAR2(256 BYTE)	Yes	30	The time zone for which the time of day is reported. Any of the longitudinal divisions of the earths surface in which a standard time is kept. Each zone observes a clock time one hour earlier than the zone immediately to the east.
LOWER_QUANTITATION_LIMIT	CHAR(12 CHAR)	Yes	31	Lower Quantitation Limit
UPPER_QUANTITATION_LIMIT	CHAR(12 CHAR)	Yes	32	upper quantitation limit
DETECTION_LIMIT	VARCHAR2(25 CHAR)	Yes	33	detection limit
DETECTION_LIMIT_DESCRIPTION	VARCHAR2(256 BYTE)	Yes	34	detection limit description
LAB_REMARK	VARCHAR2(256 BYTE)	Yes	35	Laboratory Remark

DISTANCE_MEASURE_FROM	VARCHAR2(20 CHAR)	Yes	36	Distance measured from the location.
DISTANCE_MEASURE_TO	VARCHAR2(20 CHAR)	Yes	37	Distance measured to the location.
PARTICLE_SIZE	VARCHAR2(40 CHAR)	Yes	38	User defined free text describing the particle size class for which the associated result is defined..
REPLICATE_ANALYSIS_COUNT	NUMBER	Yes	39	replicate analysis count
PRECISION	VARCHAR2(256 BYTE)	Yes	40	Estimate of the maximum possible error in the result. (e.g., Counting error in determining radiological beta particle counts.)
CONFIDENCE_LEVEL	VARCHAR2(256 BYTE)	Yes	41	confidence level
DILUTION_INDICATOR	CHAR(1 CHAR)	Yes	42	dilution indicator
RECOVERY_INDICATOR	CHAR(1 CHAR)	Yes	43	Recovery Indicator
CORRECTION_INDICATOR	CHAR(1 CHAR)	Yes	44	Correction Indicator.
STN_LATITUDE	NUMBER	Yes	45	Station Latitude measure.
STN_LONGITUDE	NUMBER	Yes	46	A statistic or calculation type which describes the reported result (e.g., average, mode, median, MPN).
STN_HDATUM	CHAR(12 CHAR)	Yes	47	Station Horizontal Datum
STN_STD_LATITUDE	NUMBER	Yes	48	Station Standard Latitude measure
STN_STD_LONGITUDE	NUMBER	Yes	49	Station Standard Longitude measure
STN_STD_HDATUM	CHAR(12 CHAR)	Yes	50	Station Standard Horizontal Datum
HYDROLOGIC_UNIT_CODE	CHAR(8 CHAR)	Yes	51	The Federal Information Processing Standard (FIPS) code for Hydrologic Units, as defined in FIPS Publication 103.
GENERATED_HUC	CHAR(8 CHAR)	Yes	52	Generated HUC. If user provided the optional Hydrologic Unit Code, the HUC is generated based on the station location information.
RESULT_IS_NUMBER	NUMBER	Yes	53	result internal serial number
ACTIVITY_MEDIUM	VARCHAR2(20 CHAR)	Yes	54	The name of the medium or matrix where the Field Activity occurred.
FK_STATION	NUMBER(12,0)	Yes	55	Foreign Key referencing DI_STATION.
FK_ORG	NUMBER(12,0)	Yes	56	Foreign Key referencing DI_ORG.
FK_DB_CAT	NUMBER(12,0)	Yes	57	Foreign Key referencing DI_DB_CAT.
FK_GEN_DB_CAT	NUMBER(12,0)	Yes	58	Foreign Key referencing DI_GEN_DB_CAT.
FK_GEO_COUNTY	NUMBER(12,0)	Yes	59	Foreign Key referencing DI_GEO_COUNTY.

FK_GEO_STATE	NUMBER(12,0)	Yes	60	Foreign Key referencing DI_GEO_STATE.
FK_DATE_ACT_START	NUMBER(12,0)	Yes	61	Foreign Key referencing DI_DATE_ACT_START.
FK_ACT_MEDIUM	NUMBER(12,0)	Yes	62	Foreign Key referencing DI_ACT_MEDIUM.
FK_ACT_MATRIX	NUMBER(12,0)	Yes	63	Foreign Key referencing DI_ACT_MATRIX.
ACTIVITY_IS_NUMBER	NUMBER(12,0)	Yes	64	Field activity internal serial number.
FK_CHAR	NUMBER(12,0)	Yes	65	Foreign Key referencing DI_CHAR.
FK_UNIT_CONVERSION	NUMBER(12,0)	Yes	66	Foreign Key referencing DI_UNIT_CONVERSION.
ACTIVITY_ID	VARCHAR2(256 BYTE)	Yes	67	Field Activity Identifier
REPLICATE_NUMBER	NUMBER(3,0)	Yes	68	User-assigned number applied to samples of the same type, medium, etc. to differentiate between them (e.g., Water Replicate 1 and Water Replicate 2, etc.) Note: This applies only to certain Activity Categories for Samples (replicate, field replicate, depletion replicate, sample created from sample) and Measurements (measurement replicate)- not Observations.
ACTIVITY_TYPE	VARCHAR2(256 BYTE)	Yes	69	Field activity type.
ACTIVITY_CATEGORY	VARCHAR2(256 BYTE)	Yes	70	A descriptor used to distinguish different kinds of samples and different kinds of Measurements and Observations. The permitted value list is controlled by the choice of TYPE_NAME (above), with the entries in TSMPRMVL for CATEGORY_TYPE_NAME_M used for activities of type Msr/Obs, while CATEGORY_TYPE_NAME_S is used for activities of type Sample.
ACTIVITY_INTENT	VARCHAR2(20 CHAR)	Yes	71	The primary reason the Field Activity occurred. Note: This is mandatory for Biological Medium, and not available for other media..
LOCATION_POINT_TYPE	CHAR(16 CHAR)	Yes	72	Location Point Type
POINT_SEQUENCE_NUMBER	NUMBER(5,0)	Yes	73	The number that indicates the sequence position of current point among a group of points. For example, if several points are recorded to mark the boundary of a station, the sequence number specifies the order in which the system should "connect the dots."
WELL_NUMBER	CHAR(15 CHAR)	Yes	74	Well Number
PIPE_NUMBER	CHAR(15 CHAR)	Yes	75	Pipe Number
ACTIVITY_STOP_D	DATE	Yes	76	Activity Stop Date time

ATE_TIME				
ACT_STOP_TIME_ZONE	VARCHAR2(5 CHAR)	Yes	77	Field activity stop time zone
ACTIVITY_REL_DEPTH	CHAR(15 CHAR)	Yes	78	Activity relative depth from the reference point
ACTIVITY_DEPTH	VARCHAR2(256 BYTE)	Yes	79	Distance from the surface to the point in the water column at which the activity is conducted.
ACTIVITY_DEPTH_UNIT	VARCHAR2(256 BYTE)	Yes	80	Field Activity depth unit
ACTIVITY_UPPER_DEPTH	VARCHAR2(256 BYTE)	Yes	81	This measure is associated with an activity that is normally conducted over or within a vertical depth range within the water column. This attribute is a measure of the distance from the surface to the upper boundary of the zone within which the activity is conducted or over which the activity is integrated.
ACTIVITY_LOWER_DEPTH	VARCHAR2(256 BYTE)	Yes	82	This measure is associated with an activity that is normally conducted over or within a vertical depth range within the water column. This attribute is a measure of the distance from the surface to the lower boundary of the zone within which the activity is conducted or over which the activity is integrated.
UPR_LWR_DEPTH_UNIT	VARCHAR2(256 BYTE)	Yes	83	Unit code for Upper and Lower depth measure
FIELD_PROCEDURE_ID	VARCHAR2(256 BYTE)	Yes	84	The abbreviated name or identifying code of the analytical procedure.
GEAR_CONFIG_ID	VARCHAR2(256 BYTE)	Yes	85	The code that identifies a Gear Configuration within the Organization. Note: Must be unique within the Organization. Used as a shorthand or abbreviation to represent the Gear Configuration in batch data update runs.
ACTIVITY_LATITUDE	NUMBER	Yes	86	Field activity latitude measure
ACTIVITY_LONGITUDE	NUMBER	Yes	87	Field activity Longitude measure
ACT_STD_LATITUDE	NUMBER	Yes	88	Field activity standard latitude measure
ACT_STD_LONGITUDE	NUMBER	Yes	89	Field activity Standard Longitude measure
ACT_STD_HDATUM	CHAR(12 CHAR)	Yes	90	field activity standard horizontal datum
STD_VALUE	NUMBER	Yes	91	Standard Value
STD_UNIT	VARCHAR2(10 CHAR)	Yes	92	Standard Unit code

FK_ACT_MAD_HDATUM	NUMBER(12,0)	Yes	93	Foreign Key referencing DI_ACT_MAD_HDATUM.
FK_ACT_MAD_HMETHOD	NUMBER(12,0)	Yes	94	Foreign Key referencing DI_ACT_MAD_HMETHOD.
ACTIVITY_ISN	NUMBER	Yes	95	Field activity internal serial number.
VISIT_START_DATE_TIME	DATE	Yes	96	Station visit start date time
VISIT_START_TIME_ZONE	CHAR(3 CHAR)	Yes	97	Station Visit Start Time Zone
VISIT_STOP_DATE_TIME	DATE	Yes	98	station visit stop date time
VISIT_STOP_TIME_ZONE	CHAR(3 CHAR)	Yes	99	Station Visit Stop Time Zone
ACTIVITY_MATRIX	VARCHAR2(256 BYTE)	Yes	100	Activity matrix
FIELD_SET	VARCHAR2(130 CHAR)	Yes	101	Field Set.
POINT_NAME	VARCHAR2(30 CHAR)	Yes	102	User-specified free text name by which a specific absolute location point will be known.
SGO_INDICATOR	CHAR(1 CHAR)	Yes	103	Station SGO Indicator
MAP_SCALE	VARCHAR2(20 CHAR)	Yes	104	map scale
FIELD_GEAR_ID	VARCHAR2(256 BYTE)	Yes	105	Field Gear ID
BIAS	VARCHAR2(256 BYTE)	Yes	106	A consistent deviation of measured values from the true value, caused by systematic errors in a procedure, as determined by applying identical procedures to a specimen of known properties.
CONF_LVL_CORR_BIAS	CHAR(1 CHAR)	Yes	107	A code indicating whether the confidence level has been corrected for Bias.
RESULT_COMMENT	VARCHAR2(4000 CHAR)	Yes	108	Result Comment
TEXT_RESULT	VARCHAR2(4000 CHAR)	Yes	109	Text result
CAS_NUMBER	VARCHAR2(500 CHAR)	Yes	110	EPA CAS number.
EPA_REG_NUMBER	VARCHAR2(500 CHAR)	Yes	111	EPA registry number.
ITIS_NUMBER	VARCHAR2(500 CHAR)	Yes	112	Integrated Taxonomy Information system number.
CONTAINER_DESC	VARCHAR2(256 BYTE)	Yes	113	sample collection container description
TEMP_PRESERVN_TYPE	VARCHAR2(25 CHAR)	Yes	114	A default for the name of the type of temperature based physical preservation.
PRESRV_STRGE_PRCDR	VARCHAR2(256 CHAR)	Yes	115	Free text description providing additional information about the handling, transport,

				preservation, and storage of the sample.
PORTABLE_DATA_LOGGER	VARCHAR2(34 CHAR)	Yes	116	Portable Data Logger
FK_STN_ACT_PT	NUMBER(12,0)	Yes	117	Foreign Key referencing DI_STN_ACT_PT.
FK_STATN_TYPES	NUMBER(12,0)	Yes	118	Foreign Key referencing DI_STATN_TYPES.
BLOB_ID	VARCHAR2(25 CHAR)	Yes	119	BLOB Identifier
BLOB_TITLE	VARCHAR2(256 BYTE)	Yes	120	User-defined Title defining or describing the BLOB Content.
ACT_BLOB_ID	VARCHAR2(25 CHAR)	Yes	121	Activity BLOB identifier
ACT_BLOB_TITLE	VARCHAR2(256 BYTE)	Yes	122	User-defined Title defining or describing the BLOB Content.
ACTIVITY_COMMENT	VARCHAR2(256 CHAR)	Yes	123	Activity comment
ACTIVITY_DEPTH_REF_POINT	VARCHAR2(256 CHAR)	Yes	124	The text that describes the reference point from which the depth is measured, typically "Surface".
PROJECT_ID	VARCHAR2(256 BYTE)	Yes	125	
TRIBAL_WATER_QUALITY_MEASURE	CHAR(1 CHAR)	Yes	126	
ACTIVITY_MEAS_QUAL_CODE	VARCHAR2(256 CHAR)	Yes	127	
ACTIVITY_COND_ORG_TEXT	VARCHAR2(256 CHAR)	Yes	128	
RESULT_DEPTH_MEAS_VALUE	VARCHAR2(256 CHAR)	Yes	129	
RESULT_DEPTH_MEAS_UNIT_CODE	VARCHAR2(256 CHAR)	Yes	130	
RESULT_DEPTH_ALT_REF_PT_TXT	VARCHAR2(256 CHAR)	Yes	131	
ANALYTICAL_METHOD_LIST_AGENCY	VARCHAR2(256 BYTE)	Yes	132	
ANALYTICAL_METHOD_LIST_VER	VARCHAR2(256 BYTE)	Yes	133	
SMPRP_TRANSPORT_STORAGE_DESC	VARCHAR2(256 BYTE)	Yes	134	
SOURCE_SYSTEM	VARCHAR2(256 CHAR)	Yes	135	
SOURCE_UID	NUMBER(20,0)	Yes	136	
ETL_ID	VARCHAR2(256 BYTE)	Yes	137	

HORIZONTAL_ACCURACY_MEASURE	VARCHAR2(256 BYTE)	Yes	138	
LAB_ACCRED_AUTHORITY	VARCHAR2(256 BYTE)	Yes	139	
METHOD_SPECIATION	VARCHAR2(256 BYTE)	Yes	140	
LAB_SAMP_PRP_METHOD_ID	VARCHAR2(256 BYTE)	Yes	141	
LAB_SAMP_PRP_METHOD_CONTEXT	VARCHAR2(256 BYTE)	Yes	142	
LAB_SAMP_PRP_METHOD_QUAL_TYPE	VARCHAR2(256 BYTE)	Yes	143	
LAB_SAMP_PRP_START_DATE_TIME	DATE	Yes	144	
LAB_SAMP_PRP_START_TMZONE	VARCHAR2(256 BYTE)	Yes	145	
LAB_SAMP_PRP_END_DATE_TIME	DATE	Yes	146	
LAB_SAMP_PRP_END_TMZONE	VARCHAR2(256 BYTE)	Yes	147	
LAB_SAMP_PRP_DILUTION_FACTOR	VARCHAR2(256 BYTE)	Yes	148	
SAMPLING_POINT_NAME	VARCHAR2(256 BYTE)	Yes	149	
LAST_CHANGE_DATE	DATE	Yes	150	

MAIN VIEWS USED TO CREATE EPA CATALOG TABLES

THE SERIESCATALOG VIEW

```
CREATE VIEW [dbo].[epa_odm_seriesCatalog] AS
SELECT     Variables.VariableName, Variables.SampleMedium,
Variables.VariableCode, Variables.variableType, Variables.DataType, SC.FK_STATION,
                      SC.FK_CHAR, SC.FK_ACT_MEDIUM, SC.ValueCount, SC.BeginDateTime, SC.EndDateTime, Sites.SiteID,
Sites.SiteName, Sites.LONGITUDE,
                      Sites.LATITUDE, Sites.County, Sites.Elevation_M, Sites.State, Sites.DESCRIPTION_TEXT, Sites.FIPS_STATE_CODE,
Sites.FIPS_COUNTY_CODE,
                      Sites.EPA_StationID, Sites.StationType, Sites.EPA_HUC, Sites.EPA_OrgID, Sites.EPA_OrgName, Sites.SiteCode,
Sites.vertical_datum,
                      'EPA' AS SiteVocabulary, 'EPA' AS VariableVocabulary
FROM       dbo.epa_disctinctSeries_no_fastfirstrow AS SC LEFT OUTER JOIN
                      dbo.epa_odm_variables AS Variables ON SC.FK_CHAR = Variables.FK_CHAR AND
                      SC.FK_ACT_MEDIUM = Variables.FK_ACT_MEDIUM LEFT OUTER JOIN
                      dbo.ODM_SITES AS Sites ON SC.FK_STATION = Sites.SiteID
```

THE VARIABLES VIEW

```
CREATE VIEW [dbo].[epa_odm_variables]
AS
SELECT     TOP (100) PERCENT dbo.DI_CHARACTERISTIC.DISPLAY_NAME AS
VariableName, dbo.DI_ACTIVITY_MEDIUM.ACTIVITY_MEDIUM AS SampleMedium,
                      CONVERT(varchar, dbo.variableDistinct.FK_CHAR) + '-' + CONVERT(varchar, dbo.variableDistinct.FK_ACT_MEDIUM)
AS VariableCode,
                      dbo.DI_CHARACTERISTIC.CAS_NUMBER AS epa_cas_number, CONVERT(varchar, NULL) AS variableType,
CONVERT(varchar,
NULL) AS DataType,
                      dbo.variableDistinct.FK_CHAR, dbo.variableDistinct.FK_ACT_MEDIUM, 'EPA' AS VariableVocabulary
FROM       dbo.variableDistinct INNER JOIN
                      dbo.DI_CHARACTERISTIC ON dbo.variableDistinct.FK_CHAR = dbo.DI_CHARACTERISTIC.PK_ISN INNER JOIN
                      dbo.DI_ACTIVITY_MEDIUM ON dbo.variableDistinct.FK_ACT_MEDIUM = dbo.DI_ACTIVITY_MEDIUM.PK_ISN
```

THE SITES VIEW

```
CREATE VIEW [dbo].[ODM_SITES]
WITH SCHEMABINDING
AS
SELECT DISTINCT
                      dbo.FA_STATION.PK_ISN AS SiteID, dbo.FA_STATION.STATION_NAME AS SiteName, dbo.FA_STATION.LATITUDE,
dbo.FA_STATION.LONGITUDE,
                      CASE WHEN isnumeric(ELEVATION) = 1 THEN Elevation ELSE NULL END AS Elevation_M,
dbo.DI_GEO_COUNTY.COUNTY_NAME AS County,
                      dbo.DI_GEO_STATE.STATE_NAME AS State, dbo.FA_STATION.DESCRIPTION_TEXT,
dbo.DI_GEO_STATE.FIPS_STATE_CODE,
                      dbo.DI_GEO_COUNTY.FIPS_COUNTY_CODE,
dbo.FA_STATION.STATION_ID AS EPA_StationID, dbo.DI_STATN_TYPES.PRIMARY_TYPE AS StationType,
                      dbo.FA_STATION.HYDROLOGIC_UNIT_CODE AS EPA_HUC, dbo.DI_ORG.ORGANIZATION_ID AS EPA_OrgID,
                      LTRIM(RTRIM(dbo.DI_ORG.ORGANIZATION_NAME)) AS EPA_OrgName,
LTRIM(RTRIM(dbo.FA_STATION.ORGANIZATION_ID))
                      + ':' + LTRIM(RTRIM(dbo.FA_STATION.STATION_ID)) AS SiteCode, NULL AS vertical_datum, 'EPA' AS SiteVocabulary,
                      dbo.FA_STATION.GENERATED_HUC AS HUC, CASE
len(GENERATED_HUC) WHEN 2 THEN CONVERT(decimal(18, 0), GENERATED_HUC +
'0000000000')
                       WHEN 4 THEN CONVERT(decimal(18, 0), GENERATED_HUC
+ '00000000') WHEN 6 THEN CONVERT(decimal(18, 0), GENERATED_HUC +
+ '000000')
                      WHEN 8 THEN CONVERT(decimal(18, 0), GENERATED_HUC
+ '0000') ELSE NULL END AS HUCNUMERIC,
                      'EPA:' +
LTRIM(RTRIM(dbo.FA_STATION.ORGANIZATION_ID)) + ':' +
LTRIM(RTRIM(dbo.FA_STATION.STATION_ID)) AS ws_siteCode
FROM       dbo.FA_STATION LEFT OUTER JOIN
                      dbo.DI_STATN_TYPES ON dbo.FA_STATION.FK_STATN_TYPES = dbo.DI_STATN_TYPES.PK_ISN LEFT OUTER JOIN
                      dbo.DI_GEO_COUNTY ON dbo.FA_STATION.FK_GEO_COUNTY = dbo.DI_GEO_COUNTY.PK_ISN LEFT OUTER JOIN
                      dbo.DI_ORG ON dbo.FA_STATION.FK_ORG = dbo.DI_ORG.PK_ISN LEFT OUTER JOIN
                      dbo.DI_GEO_STATE ON dbo.FA_STATION.FK_GEO_STATE = dbo.DI_GEO_STATE.PK_ISN
WHERE     (dbo.FA_STATION.FK_ORG IS NOT NULL)
```

APPENDIX C. CURRENT TAGGING STATISTICS (APRIL 6, 2010)

NetworkName	NetworkTitle	variables	Searchable_vararriables	series	searchable_serie	values	Searchable_values	Var%	series %	values %
NWISDV	NWIS Daily Values	246	157	51797	50706	274762525	270847574	**63.82%**	**97.89%**	**98.58%**
NWISIID	NWIS Instantaneous Irregular Data	605	202	1635298	957402	15501245	11667688	**33.39%**	**58.55%**	**75.27%**
NWISUV	NWIS Unit Values	128	3	26188	7326	77935488	21802176	**2.34%**	**27.97%**	**27.97%**
EPA	EPA STORET	3032	1143	4849161	3204442	94349967	64151710	**37.70%**	**66.08%**	**67.99%**
NWISGW	NWIS Ground Water	1	1	827200	827200	8325633	8325633	**100.00%**	**100.00%**	**100.00%**
NCDCISH	NCDC Hourly Data	35	3	866950	74310	0	0	**8.57%**	**8.57%**	**0.00%**
RiverGages	USACE River Gages	48	7	3077	192	263101295	16417163	**14.58%**	**6.24%**	**6.24%**
CIMS	Chesapeake Bay Information Management System	80	63	15123	12016	5418382	4624477	**78.75%**	**79.46%**	**85.35%**
MudLake	Mud Lake, Idaho, USA	26	23	50	44	1242931	1090020	**88.46%**	**88.00%**	**87.70%**
SDRPF	San Diego River Park Foundation	8	7	132	114	2424	2111	**87.50%**	**86.36%**	**87.09%**
COTCsnow	Crown of the Continent Observatory Snow	5	5	10	10	37734	37734	**100.00%**	**100.00%**	**100.00%**
SFe_Storet	Santa Fe, STORET	5	5	103	103	4484	4484	**100.00%**	**100.00%**	**100.00%**
SFe_MICROWAVECITRA	Santa Fe MICROWAVECITRA	4	4	8	8	219116	219116	**100.00%**	**100.00%**	**100.00%**
NADP	National Atmospheric Deposition Program	13	12	247	228	1713885	1577793	**92.31%**	**92.31%**	**92.06%**
SRBHOS	Susquehanna River Basin Hydrologic Observatory	33	32	35	35	2182460	2182460	**96.97%**	**100.00%**	**100.00%**
TIFP_LowerSabine	Texas Instream Flow, Lower Sabine	17	3	970	185	970	185	**17.65%**	**19.07%**	**19.07%**
LittleBearRiver	Little Bear River Experimental Watershed, Northern Utah, USA	58	50	266	227	5015598	4558771	**86.21%**	**85.34%**	**90.89%**
RCEW	Reynolds Creek Experimental Watershed	60	58	373	369	17816604	17391266	**96.67%**	**98.93%**	**97.61%**
IIHRTippingB	IIHR Tipping Bucket	1	1	3	3	275017	275017	**100.00%**	**100.00%**	**100.00%**
IIHRNexrad	IIHR Nexrad	1	1	142	142	3356312	3356312	**100.00%**	**100.00%**	**100.00%**
IIHRWQ	IIHR Water Quality	4	4	11	11	84631	84631	**100.00%**	**100.00%**	**100.00%**
BaltimoreGW	Baltimore Waters Test Bed Ground Water Level Data	2	1	28	14	333045	169795	**50.00%**	**50.00%**	**50.98%**

SFe_CTDSondes	Santa Fe CTD Sondes	1	1	6	6	262347	262347	**100.00 %**	**100.00 %**	**100.00 %**
SantaFe-RainDlySRWMD	Santa Fe Daily Rain Tipping Bucket	1	1	11	11	14371	14371	**100.00 %**	**100.00 %**	**100.00 %**
MPE	Multi-sensor Precipitation Estimates	1	1	9308	9308	5010626	5010626	**100.00 %**	**100.00 %**	**100.00 %**
SBRP	Superfund Basic Research Program	47	4	1504	128	3008	256	**8.51%**	**8.51%**	**8.51%**
MAST	MAST	4	4	84	84	25480	25480	**100.00 %**	**100.00 %**	**100.00 %**
Sonora	Sonora, Mexico project	5	1	20	4	22	6	**20.00 %**	**20.00 %**	**27.27 %**
TWDB_Sondes	TWDB_Sondes	6	5	364	292	182697	146986	**83.33 %**	**80.22 %**	**80.45 %**
TWDB_Wind	TWDB Wind	2	2	164	164	86408	86408	**100.00 %**	**100.00 %**	**100.00 %**
LBL	Grasslands Ecological Area of the San Joaquin Basin, California	6	5	110	108	6439473	6383527	**83.33 %**	**98.18 %**	**99.13 %**
HydroNEXRAD	HydroNEXRAD	1	1	215355	215355	3773019600	3773019600	**100.00 %**	**100.00 %**	**100.00 %**
BaltPrecip	Baltimore Precipitation	1	1	5	5	34238	34238	**100.00 %**	**100.00 %**	**100.00 %**
ODMDCEW2	Dry Creek Experimental Watershed, SW Idaho	20	19	254	240	22633716	22563345	**95.00 %**	**94.49 %**	**99.69 %**
ParadiseCreek	Paradise Creek Watershed, Idaho	12	12	37	37	213267	213267	**100.00 %**	**100.00 %**	**100.00 %**
LoganRiver	Logan River Observations, Northern Utah, USA	5	5	18	18	877985	877985	**100.00 %**	**100.00 %**	**100.00 %**
NPCA	Niagara Peninsula Conservation Authority Water Data Service	9	4	32	14	666654	145211	**44.44 %**	**43.75 %**	**21.78 %**
RIMP	RIM Program	13	6	108	54	153070	60522	**46.15 %**	**50.00 %**	**39.54 %**
BEACON_IBM	Beacon Institute for River and Esturay	57	4	122	11	1910825	348315	**7.02%**	**9.02%**	**18.23 %**
	Totals:	*4603*	*1861*	*8504674*	*5360926*	*4583213533*	*4237978606*	***40.43 %***	***63.04 %***	***92.47 %***