

Developing and Integrating Web Services for Georeferencing to the NHD

James Rineer, RTI International (RTI)

Anne Marie Miller, RTI

James Sinnott, RTI

Michael Plastino, U.S. Environmental Protection Agency (EPA), Office of Water, Immediate Office

RTI International (RTI), under a contract with the U.S. Environmental Protection Agency (EPA), has developed a Web-based ArcIMS application, called the Web-Based Reach Indexing Tool (WebRIT). WebRIT is used to create, validate, and update National Hydrography Dataset (NHD) reach indexes, latitude/longitude coordinates, and other locational data. The initial version of WebRIT used a traditional ArcIMS architecture. However, recent advances in Web service technology allow WebRIT to take advantage of existing Web services and build new ones for use by other EPA applications. This paper describes the WebRIT architecture, Web services created and used by WebRIT, and the advantages of the application architecture.

Background

WebRIT for Watershed Assessment, Tracking, and Environmental Results (WATERS) is an Internet mapping tool that allows users to submit and ultimately update locational data for a wide variety of environmental programs. To help users identify locations, WebRIT provides point-and-click tools and displays data from various EPA environmental programs, as well U.S. Geological Survey (USGS) topographical maps and aerial photos, and the NHD. By assisting users in georeferencing their environmental program data to the NHD, the WebRIT greatly enhances EPA's ability to evaluate the success of programs and target optimal environmental protection strategies.

Currently, WebRIT allows users to submit locational information for the Beaches Environmental Assessment and Coastal Health (BEACH) Act, the Nonpoint Source Grant Reporting and Tracking System (GRTS) program, the National Pollutant Discharge Elimination System (NPDES) permit program, and USGS's National Water Information System (NWIS) program. In future versions, the WebRIT will be available for other program data, such as data from the STorage and RETrieval (STORET) database and other water quality monitoring stations, 303(d) impaired water listings, and more.

The NHD (<http://nhd.usgs.gov/>) is a nationally consistent spatial data set of U.S. surface waters that contains information for several types of surface water features, including streams, lakes, reservoirs, canals, and dams. It is the result of cooperative efforts between EPA and the USGS. Every surface water feature in NHD is represented by a unique identifier (reach code). The benefit of georeferencing environmental program data to NHD is that the information becomes anchored to a nationally consistent identifier (reach code). This provides a framework for integrating many national surface water and environmental program databases. The NHD currently provides the framework for EPA's WATERS data integration effort (<http://www.epa.gov/waters>).

The WebRIT assigns Entity IDs or attributes to user selected portions of NHD through the creation of "events" that are stored in a central EPA database. These events represent the location of an attribute (such as the streams affected by a nonpoint source grant) along NHD stream reaches. Event records are stored in a table that contains the NHD reach code (unique stream reach address), position(s) along the NHD reach, and a unique identifier that links to the program database (such as the identifier for a specific impaired water). These events are what

the user creates when using the WebRIT. The WebRIT allows the user to create three types of event layers: points, lines, and polygons. The WebRIT also provides methods for automated NHD indexing (event creation). The events created through the WebRIT interface are stored in a central EPA database where they can be later reviewed, modified, and approved. After approval, the events stored in the WebRIT database are transferred to the NHD Reach Address Database (RAD) (<http://www.epa.gov/waters/about/rad.html>), where they can be viewed through other EPA tools (<http://www.epa.gov/waters/tools/>).

The WebRIT contains many datasets from both EPA and USGS. Some of the datasets include: NHD, Digital Raster Graphics (DRGs), Digital Orthophoto Quadrangles (DOQs), detailed streets, county boundaries, zip codes, etc. The WebRIT not only allows users to create NHD events, but also has the ability to print out topographic maps and aerial photographs with user-defined points displayed on top. The user can also add annotation (an identifier and description) associated with each of these user-defined points. The WebRIT also provides users the ability to capture latitude/longitude coordinates, many spatial/attribute selection options, and the ability to capture Federal Geographic Data Committee (FGDC) compliant metadata associated with the event data created. Future versions of the WebRIT will also provide other latitude-longitude services and the ability to update existing environmental program data that have already been submitted to EPA.

This paper discusses the WebRIT architecture and design, WebRIT's current and future use of Web services, and WebRIT's automated indexing approach. For additional information on georeferencing, browse to <http://www.epa.gov/owow/monitoring/georef/>. For additional information on WebRIT, browse to <http://www.epa.gov/waters/webrit/>. For additional information on NHD, browse to <http://nhd.usgs.gov/>.

WebRIT Architecture

The WebRIT utilizes a distributed application architecture that breaks the application into smaller logical parts to handle presentation functions, business logic, and database functionality. This approach is a popular choice for Internet application development because it makes the code easier to write, easier to maintain, and easier to reuse.

Physically, the application can be broken into three tiers. The three tiers are the interface tier, the middle tier, and the database tier. The interface tier is composed of the client's browser and code that is run inside of it including HTML, DHTML, and JavaScript. The middle tier represents the Web server, server side code, and server extensions including Active Server Pages (ASP), Custom ActiveX Dynamic Link Libraries (DLLs), ActiveX Data Objects (ADO), Oracle Objects for OLE (OO4O), Servlets, and ArcIMS. The database tier includes the database, Oracle, Oracle PL/SQL stored procedures, and the database extensions including Oracle Spatial and the Environmental Systems Research Institute's (ESRI's) Spatial Data Engine (SDE). The database tier contains both spatial and nonspatial tables. Figure 1 depicts the physical layout of the WebRIT architecture.

WebRIT Architecture Overview

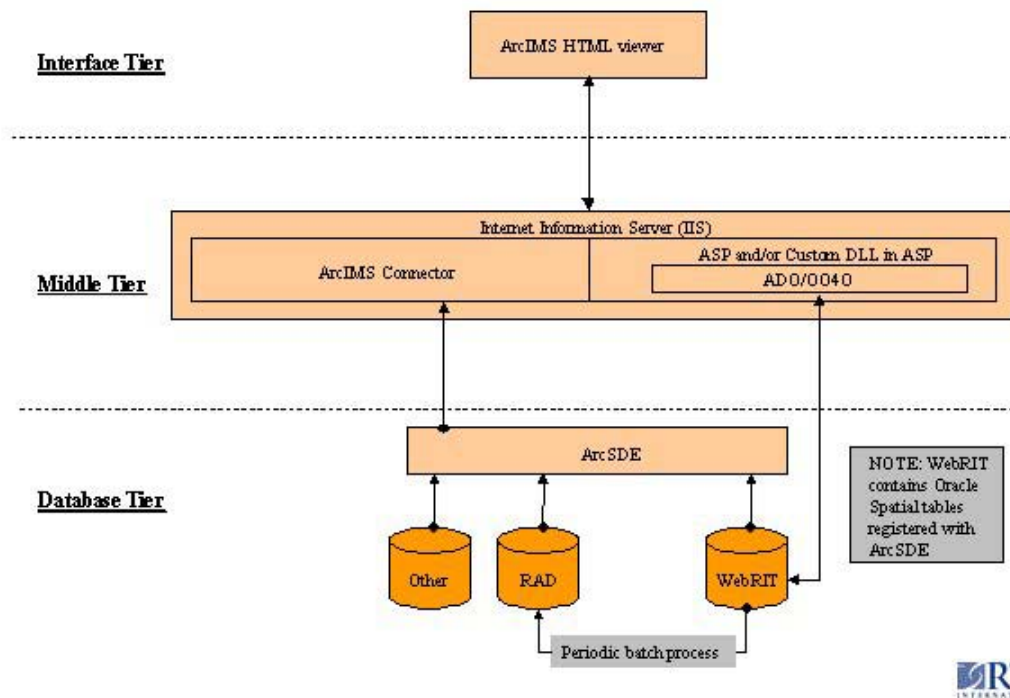


Figure 1: WebRIT Architecture Overview

The decision to place code in one tier versus another is made by looking at the advantages that each of the tiers may have over the others. Code may be placed in a specific tier to enhance performance, take advantage of the tier's unique programming benefits, or leverage code for future reuse, perhaps as a Web service. There is overlap of presentation functionality, business logic, and database functionality among the three physical tiers. For example, the client's browser with JavaScript, HTML, and DHTML primarily handles the presentation functionality but ASP in the middle tier also plays a role in creating dynamic content for some pages. In addition, business logic is split between the middle and database tiers. ASP with OO4O is used to execute dynamically generated SQL statements, but stored procedures are also used to increase performance and make the code easier to write and maintain.

The WebRIT makes extensive use of procedures stored in the database. Using stored procedures moves the application's business logic closer to the database to increase efficiency and also to take advantage of PL/SQL and Oracle Spatial SQL calls. In addition, the stored procedures can be easily bundled into future Web services providing some WebRIT functionality without the existing interface or middle tiers. Consumers of WebRIT services can create their own interface and middle tiers.

WebRIT Design

The WebRIT application is based on the ArcIMS HTML viewer model utilizing only HTML, DHTML, JavaScript, and frames in the client's browser. Frames are used to organize the interface features into groups. The primary frames are the toolbar, table of contents, map window, and information window. The use of frames makes the coding more modular and thus easier to maintain and enhance without complete rewrites. A potential drawback to this method

is that it introduces many different languages to the coding effort. For example, to code for the selection tool, the developer needs to understand HTML, DHTML, JavaScript, ASP, Visual Basic, SQL, PL/SQL, and Oracle Spatial syntax. But a big advantage of the design is that any one of these pieces can be modified, upgraded, or replaced without rewriting the code for the entire tool. Figure 2 shows the interface of the WebRIT application.

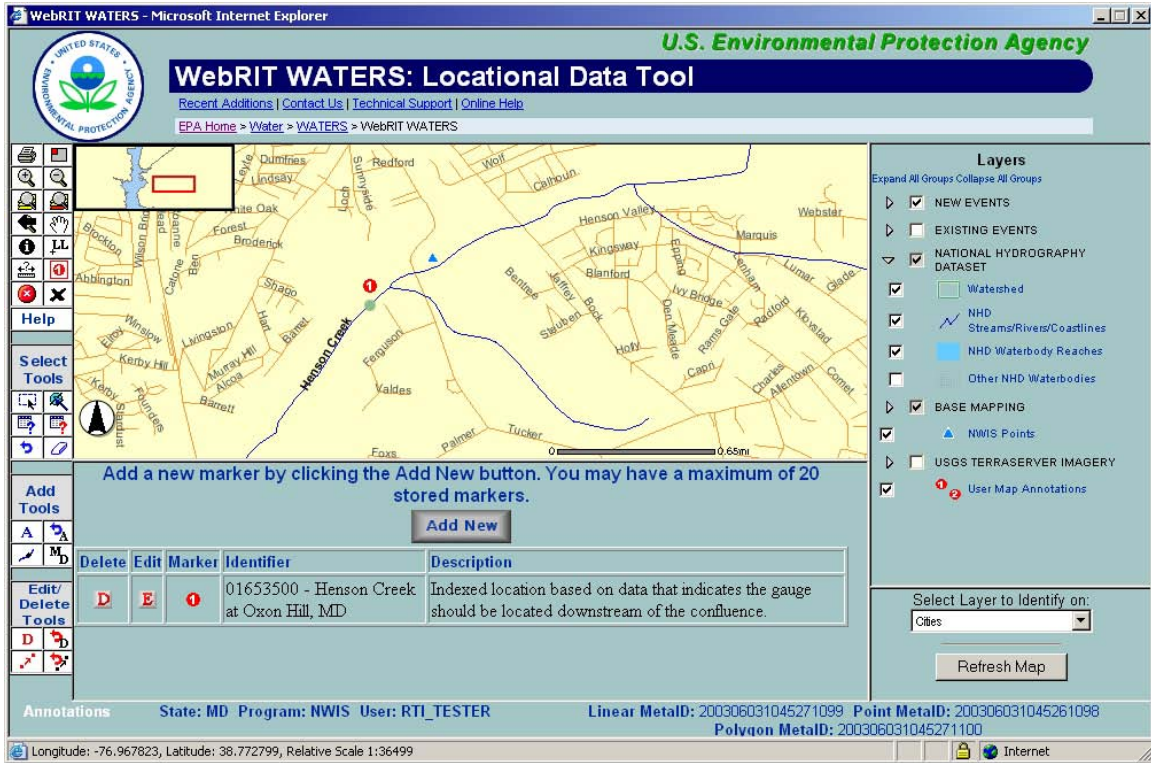


Figure 2: WebRIT Interface

Figure 2 shows an example of the results of indexing an USGS NWIS stream gage to the NHD. The blue triangle represents the un-indexed location for the gage. The small green point on the NHD below the confluence represents the user-indexed location for this gage. In this example, based on other data, the user knows that the gage should fall just below the confluence of the two streams. An annotation was added by the user for this location and is indicated by the red circle with the number 1 in it. The number corresponds to the table of annotations in the information frame, where the user indicates that indexed location was chosen based on other data that indicated the gage should be downstream of the confluence. The current map view with the annotation, un-indexed gage location, and indexed gage location can then be printed for the user's records or review. The new indexed gage is kept in the WebRIT tables until it is approved and then it is moved in a periodic batch process into the final RAD database where other tools can view it.

Advanced Features

The WebRIT application adds a number of advanced features to the out-of-the-box HTML viewer. Some of these advanced features are highlighted in this paper, including the collection of new spatial data, review and editing of existing spatial data, a user specific annotations layer, an enhanced table of contents, and the incorporation of Web services.

Collection of new spatial data is the primary goal of the WebRIT. In the current versions of ArcIMS, addition or modification of spatial data is not easily accomplished and the ability to perform linear referencing functions is not supported. WebRIT requires linear referencing to create event data. Figure 3 is an example of indexing a new linear feature to the NHD. The red stream represents a 303(d) existing impaired stream. The light green stream is the NHD indexed stream the user created with the WebRIT.

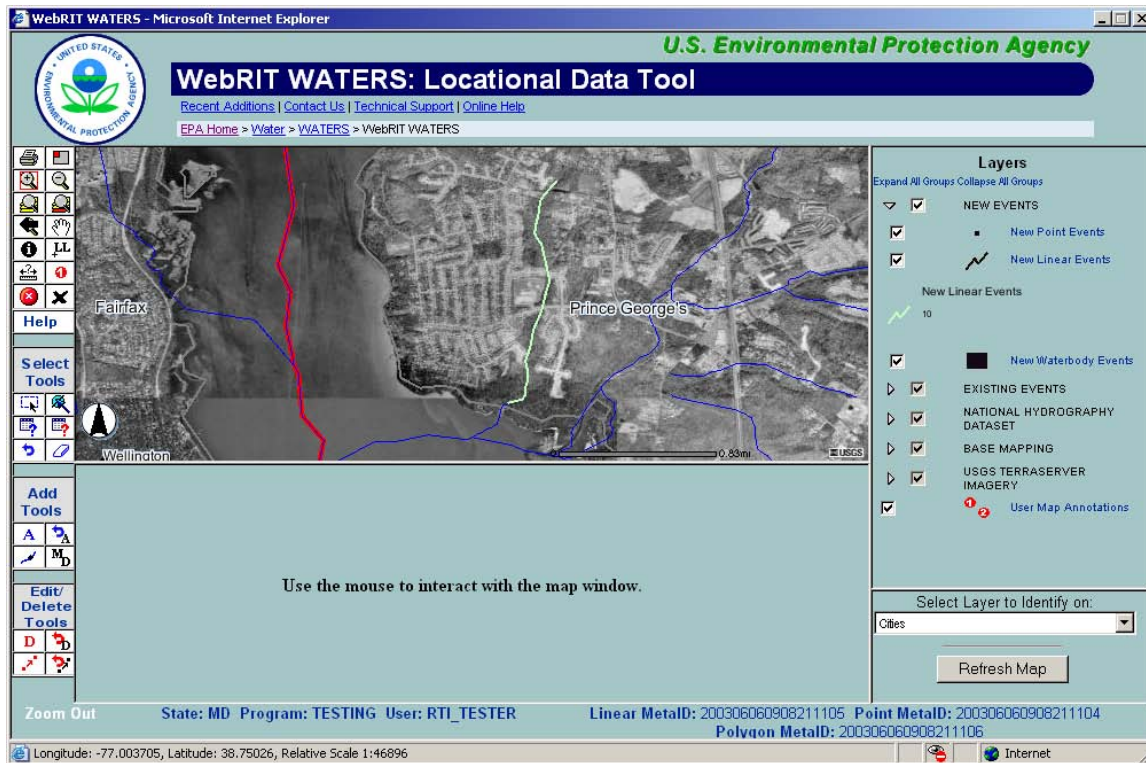


Figure 3: Linear Indexing

With the addition of Oracle Spatial to the existing Oracle SDE infrastructure, the WebRIT is able to provide tools to analyze and manipulate spatial data and perform linear referencing functions in an efficient manner. The WebRIT calls stored procedures written in PL/SQL that contain Oracle Spatial commands. Changes in spatial tables that are registered through SDE are visualized back in ArcIMS in a timely fashion. The complete loop is shown in Figure 1. The process starts in the interface tier and moves downward through the middle tier to the database tier where spatial data are added or changed via stored procedures. The process then moves upward back to the interface tier through SDE and the middle tier ArcIMS map services. Using stored procedures with Oracle Spatial commands allows the creation of a library of functions for the collection and manipulation of spatial data. One procedure can be enhanced or swapped out without breaking the entire application and the procedures can be reformed into valuable Web services for other developers wishing to add indexing to their applications.

A user-specific annotations layer was added to the application. This was done by creating a nonspatial annotations table in Oracle and allowing users to add up to 20 annotations records a piece. The application stores the annotation's location as *x,y* fields in the table along with the annotation text. Then, similar to concept of an event theme in ArcView, the table is queried for users with matching log on identifiers and annotation markers are added to the map on the acetate layer. The user is presented with a table in the information frame containing the annotation text as well as buttons to create, edit, or delete existing annotations. The user's annotation layer is listed in the WebRIT table of contents and can be turned on/off similar to any

other layer in the application. When the user prints the map, using the print tool provided, the user's specific annotation marker, identifier, and description are displayed on the map. An example of the annotation functionality can be seen in Figure 2.

Another function that will eventually be added to the WebRIT is the ability to collect features of varying geometry that are not necessarily indexed to the NHD. For example, this functionality could be used to collect or correct the locations of underground injection control (UIC) wells or facility locations.

Web Services and Future Functionality

The WebRIT is currently being enhanced to utilize existing Web services, such as Microsoft's TerraServer. Web services are also being designed to expose some of WebRIT's reach indexing (linear referencing to NHD) functionality to other applications.

The WebRIT application is designed to be able to use existing Web services easily. The Microsoft TerraServer (<http://terraserver-usa.com>) is a service that makes imagery from the USGS available to the public. The two types of imagery that are available include DOQs and DRGs.

The TerraServer supports the OpenGIS Web Map Server specification version 1.1.1, which defines standard operations to retrieve maps over the Internet. The WMS specification provides three operations that are called from a URL—GetCapabilities, GetMap, and GetFeatureInfo. GetFeatureInfo is optional and not used with the TerraServer service. GetCapabilities is called to return information on the service itself. The response to the GetCapabilities operation is a XML document that describes the content of the service and request parameters for the GetMap and GetFeatureInfo operations. By using the documented parameters from GetCapabilities, it is possible to build dynamic mapping request URLs by swapping out the type and/or coordinate box variables.

Utilizing DHTML and JavaScript, the WebRIT adds the USGS imagery from the Microsoft TerraServer service as layers to the application. Images from ArcIMS are layered on top of the imagery in a seamless manner. The WebRIT table of contents was enhanced to include the ability to list the USGS imagery as layers. See Figure 4 for an example of the use of the TerraServer imagery and the enhanced table of contents.

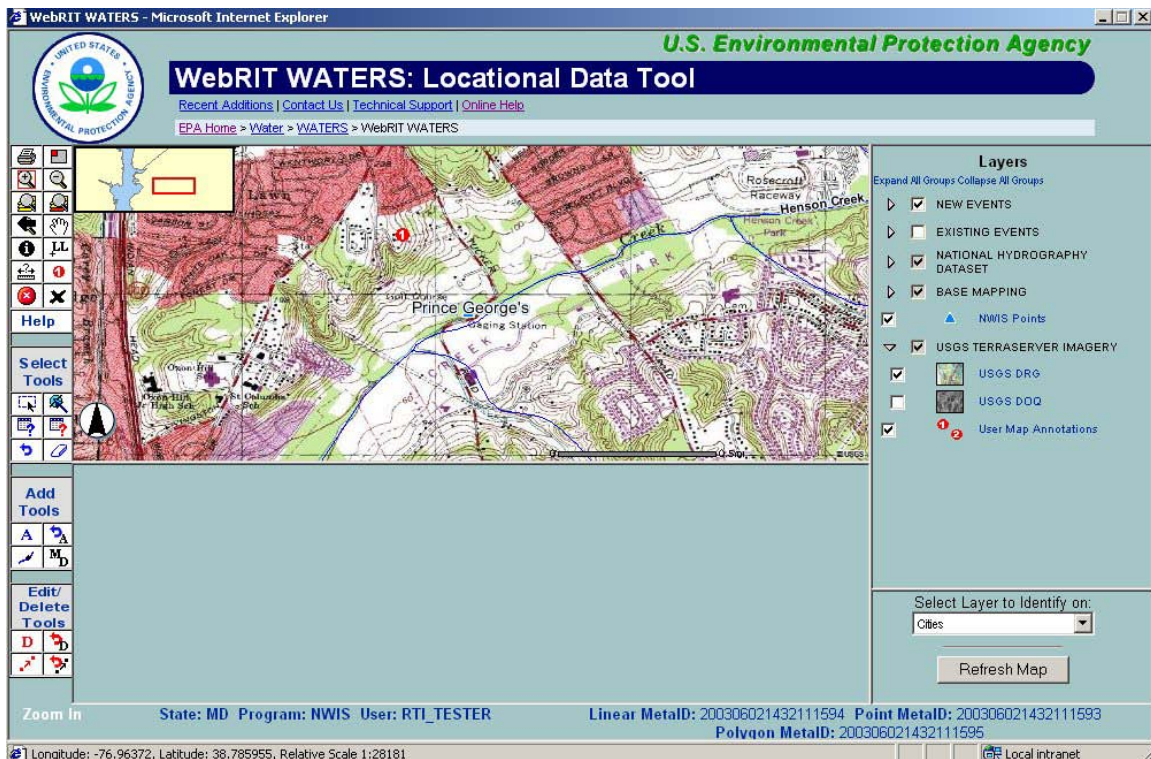


Figure 4: TerraServer USGS Imagery

The EPA is also currently designing a Web service for navigating the NHD. Future releases of the WebRIT will call this service to facilitate indexing data by allowing users to select all NHD features upstream or downstream from a user supplied location or all NHD features between two user-specified locations.

In addition to consuming Web services, future releases of the WebRIT will likely publish Web services. These Web services will allow consumers to locate NHD features close to user specified locations and even create point, line, and areal events from almost any calling environment, bypassing the WebRIT interface altogether. For instance, in the future, a user who wanted to define the area covered by a specific grant as upstream of a certain location may be able to accomplish this through a series of calls to Web services:

- Invoke a WebRIT Web service to determine the NHD feature closest to the user-supplied location
- Invoke a navigation Web service to determine the NHD features upstream from this location
- Invoke a WebRIT Web service to create events representing these NHD features
- Invoke an image service to visualize the events that were created for this request.

Currently, the existing WebRIT database tables and Oracle procedures that perform the basic indexing tasks are also being leveraged to perform much larger, automated indexing tasks. Procedures have already been written to allow batch indexing of point data. These procedures read user input locations and associated metadata out of existing and new WebRIT tables and write output to existing WebRIT event tables. These batch procedures have already been used to index more than 9,000 combined sewer overflow outfall locations and more than 112,000 pipe outfall locations. The current batch indexing procedures could be extended to handle batch indexing of linear or areal features and are also candidates for future Web services.

Conclusion

The WebRIT distributed application architecture and design makes the code easier to write, easier to maintain, and easier to reuse. It is possible to add a new EPA program quickly to the application to index new features or to add batched indexed features to the application for programs to review and correct. In addition, the architecture makes it is easy to consume Web services like the TerraServer service and future Web services. Also, in the future, components of the WebRIT may be provided as services to other developers to include in their own interfaces. The components of the WebRIT form a starting point for a standard way of collecting, reviewing, and correcting spatial data for EPA and others.

Acknowledgments

The work described in this paper was funded by the U.S. EPA under Contract 68-C-01-001 with RTI. RTI gratefully acknowledges this support.

Disclaimer: Although the research described has been funded wholly or in part by the U.S. EPA Contract 68-C-01-001 to RTI, it has not been subject to the Agency's review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

RTI International. September 2001. *NHD Reach Indexing Tool (NHD-RIT) Data Structure for Events, Metadata, and Transaction Tables*.

RTI International. October 2002. *Web-Based Reach Indexing Tool (WebRIT), Version 1.0, Design Document*.

RTI International. October 2002. *Web-Based Reach Indexing Tool (WebRIT), Version 1.0, Physical Design Document*.

RTI International. March 2003. *Web-Based Reach Indexing Tool (WebRIT) Requirements*.

Author Information

James Rineer, PE
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709
jrin@rti.org

Anne Marie Miller
RTI International
3040 Cornwallis Road
RTP, NC 27709
ammiller@rti.org

James Sinnott
RTI International
3040 Cornwallis Road`
RTP, NC 27709
jsinnott@rti.org

Michael Plastino
US EPA Office of Water, Immediate Office
EPA East
1200 Pennsylvania Avenue NW (4102M)
Washington, DC 20460
Plastino.Michael@epamail.epa.gov