CDI SSF CATEGORY 3: DATA AND INFORMATION ASSETS

USGS PROTOCOL LIBRARY: AN IMPLEMENTATION BASED ON THE NATIONAL ENVIRONMENTAL METHODS INDEX

APPLICANTS/PRINCIPAL INVESTIGATORS:

Daniel J. Sullivan, USGS Wisconsin Water Science Center, 8505 Research Way, Middleton, WI 53562. Ph (608) 821-3869 Email djsulliv@usgs.gov

Jacquelyn L. Schei, Pacific Northwest Aquatic Monitoring Partnership, Cook, WA. Ph 509.538.2299 x282 Email jschei@usgs.gov

Kathryn M. Schoephoester, USGS Center for Data Analytics, 8505 Research Way, Middleton, WI 53562. Ph (608) 821-3892, kmschoep@usgs.gov

ABSTRACT:

Develop an implementation of a protocol library using the existing NEMI database. Prototype will contain examples of major types of protocols identified in PNAMP and Monitoringmethods.org databases.

TOTAL FUNDING AMOUNT REQUESTED: $50,087

TOTAL IN-KIND FUNDING: $27,733

DATASETS:
NEMI, PNAMP

GEOGRAPHIC/GEOLOGIC/ECOSYSTEM/HABITAT/TAXONOMIC/OTHER CONTEXT:
US national, international

TYPE OF PRODUCT(S) GENERATED:
Website (extension of existing http://www.nemi.gov/)

SUMMARY
Introduction

A number of monitoring method and protocol libraries exist. Although these systems have been tailored to certain disciplines or research foci, the underlying principles, mechanisms, and processes of the systems have commonalities that could facilitate synthesizing content and information. In 2012, several of the major USGS monitoring method and protocol libraries were identified and characterized (Schei and others, CDI report, 2012).

The goals of the project that was completed in 2012 were stated as "to initiate actions that increase awareness of the existence of method and protocol libraries, promote input of content into the systems, and identify commonalities that may lead to promoting interoperability, increasing efficiencies, and minimizing redundancy." This proposal is designed to take the work in 2012 and to present a working repository of protocol information, and to create the capability for additional protocols to be added to the database at any time.

Background

The National Environmental Methods Index (NEMI, http://www.nemi.gov/) is a web-based tool that primarily contains summaries for chemical, microbiological, biological, toxicity, physical and regulatory methods. NEMI has been publicly available since 2002 and since then has seen steady user growth that includes international users as well as a core user group based in the U.S.

NEMI was created by the Methods and Data Comparability Board (MDCB), a subgroup within the National Water Quality Monitoring Council (NWQMC). The MDCB and the NWQMC are co-chaired by the U.S. Geological Survey (USGS) and the U.S. Environmental Protection Agency (EPA). The MDCB is a partnership of water quality experts from federal agencies, states, municipalities, industry, and private organizations. Both the MBCD and the NWQMC are subgroups of the Advisory Committee on Water Information (ACWI), chartered under the Federal Advisory Committee Act (FACA). The subgroups were formed in 1997 to develop a voluntary, integrated, and nationwide water quality monitoring strategy. The MDCB and the NWQMC were tasked with the review and evaluation of national water quality monitoring activities and the development of recommendations for improvements.

The NEMI system was launched to help scientists document and compare methods so they could select the most appropriate method for their specific monitoring program. NEMI was identified as a high priority of the MDCB to help support monitoring program planning by providing a place for users to compare methods. Initial funding for developing NEMI came from the USGS and the EPA and the original focus was on methods related to water quality. Within a year or so of release of the tool, methods for other media types (soil/sediment, air, animal tissue, etc.) were added into the database. Presently there are over 1,100 methods in the database and it continues to grow steadily. Recently statistical information and water-quality sensors have been added.

During the first quarter of fiscal year 2013, an updated version of NEMI will be launched that is a natural evolution of previous version while providing increased robustness, improved user interface, and enhanced searching capabilities. Currently a user questionnaire is being circulated to solicit feedback from NEMI users.

CDI SSF Category

Data and Information Assets

Project Title

USGS Protocol Library: An Implementation Based on the National Environmental Methods Index

Personnel

Dan Sullivan
Supervisory Hydrologist
USGS Wisconsin Water Science Center
8505 Research Way
Middleton, WI
608.821.3869
djsulliv@usgs.gov

Jacquelyn L. Schei
Pacific Northwest Aquatic Monitoring Partnership
Cook, WA
509.538.2299 x282
jschei@usgs.gov

Cheryl Buchwald
Hydrologist
NWQMC Methods and Data Comparability Board Executive Secretary
USGS Wisconsin Water Science Center
8505 Research Way
Middleton, WI
608.821.3873
cabuchwa@usgs.gov

Kathryn Schoephoester
Database Developer
USGS Center for Integrated Data Analytics
8505 Research Way
Middleton, WI
608.821.3892
kmschoep@usgs.gov

Mary Bucknell
Application Developer
USGS Center for Integrated Data Analytics
8505 Research Way
Middleton, WI
608.821.3920
mbucknell@usgs.gov

Geographic/geologic/ecosystem/habitat/taxonomic/other context

NEMI has users and content providers from all over the United States and users from all over the world.

SCOPE

The proposed project consists of modifying and thus extending the capabilities of the existing NEMI methods compendium. To accomplish the task of incorporating a broad array of protocols, NEMI developers will conduct user requirement queries of protocol owners. In addition, protocol owners will be asked how they would prefer to access the data. Input forms would also be created to accommodate the desires of protocol owners as much as possible.

If this project is funded, it will be another step toward a common framework for storage and retrieval of methods and protocol information. Use of this existing and proven database will greatly enhance the accessibility of this information and will provide researchers the opportunity to track versions of protocols as they evolve over time.

TECHNICAL APPROACH

The NEMI data are stored in an Oracle relational database in a simple data model that stores WHAT a user is trying to detect (chemical analyte, species, etc.), HOW a user can detect it (method, protocol, etc.), and the EQUIPMENT necessary (instrumentation, etc.). To store protocols in this data model would involve determining high-value fields that users would like to see, mapping them to existing tables and columns in the NEMI data model and adding columns and/or tables if necessary to store additional information about a protocol. This has already been done to accommodate other non-laboratory method types, such as statistical methods. The simplicity of the data model allows for much flexibility of data storage without violating the core model and relationships.

Existing protocols would be batch-loaded from the existing Access database to the Oracle database through an extract-transform-load process. Future protocols would be entered by protocol owners through data entry forms. There are currently data entry forms for entering "traditional" laboratory methods and statistical methods. Protocols would be entered by external users into existing data entry forms that would be modified to expose appropriate fields relevant to protocol data. There is an existing process that ensures a QA/QC approval of methods before they are served out to the public in final form that the protocol data entry process would adhere to by default.

Protocols would be queried through the NEMI interface just like existing methods are queried, e.g., keyword search, by title and/or species. Currently in NEMI, a summary view of methods is displayed to show a few high-value fields on one page. A summary view of protocol information would be handled in the same way, with protocol-specific fields being displayed in a summary view.

PROJECT EXPERIENCE

The NEMI team has over 10 years of experience in the development of a methods database. Dan Sullivan was the original developer of NEMI, and will be the project manager of the proposed project. Dan also has extensive field experience as a hydrologist, and has helped develop sampling protocols. Cheryl is an experienced database manager who maintains the water use database for the US Geological Survey and has strong organizational skills. Kathy Schoephoester of the Center for Integrated Data Analytics (CIDA) is currently the technical lead for NEMI, a position she has held for over five years. Kathy has seven years' experience developing relational databases in Oracle and developing web-based data entry forms and search tools. Mary Bucknell is an experienced object-oriented programmer with over 20 years' experience in developing web interfaces. Mary has been extensively involved with the redesign of the NEMI interface. Jacque Schei of PNAMP is a protocol expert and will provide invaluable input on

important protocol information to include in the proposed NEMI extension as well as a conduit for communications with the CDI Protocol Library team.

COMMITMENT TO EFFORT

This is one of the strongest aspects of this proposal. NEMI has been in existence since 2000, and publicly available since 2002. Its user base has grown steadily over the years and the financial commitment of the USGS Office of Water Quality and EPA's Office of Water has been equally steady. NEMI continues to evolve on all fronts. In 2013, an effort to link NEMI methods information with USGS' water-quality database are planned that will further enhance the usefulness of the data.

BUDGET

Budget Category	Federal Funding "Requested"	Matching Funds "Proposed"
1. SALARIES:		
Personnel		
Kathy Schoephoester, IT Specialist	$8,850	$0
Mary Bucknell, IT Specialist	$17,280	$0
Cheryl Buchwald, Data Management	$4,650	$0
Jacqu Schei, Biologist		$2,000
Total Salaries:	$30,780	$0
2. FRINGE BENEFITS:		
Personnel		
State Personnel (51%)	$0	$0
Total Fringe Benefits:	$0	$0
3. FIELD EXPENSES:		
Annual CDI Meeting		
Per diem	$330	$0
Lodging	$745	$0
Airfare	$600	$0
Vehicle Cost	$450	$0
Total Travel Expenses:	$2,475	$0
4. OTHER DIRECT COSTS: (itemize)		
CIDA Infrastructure	$0	$12,881
Operations and Administration	$0	$12,852
Wisconsin WSC General Overhead	$16,832	$0
	$0	$0
	$0	$0
	$0	$0
Total Other Direct Costs:	$16,832	$25,733
Total Direct Costs:	$30,780	$0
GRAND TOTAL:	$50,087	$27,733

TIMELINE

User requirements gathering will commence immediately upon the award of funding, and will be completed by June 1, 2013. Once the requirements are well-understood, the data model will be modified as necessary to accommodate those requirements. The updated data model will be complete by July 1, 2013. After the data model is finalized, data entry forms and the query interface will be developed to efficiently input protocols and serve out the protocol information. There will be ongoing user testing and acceptance testing in all phases of development, so some of these tasks may overlap in time.

By the end of August, 2013, a number of protocols from as many media as possible will be entered and available for query in NEMI. This will be accomplished by the existing NEMI development team based in Middleton, Wisconsin, with input from interested members of the CDI Protocol Library group from 2012, as well as the NEMI steering committee members of the NWQMC. Existing USGS and EPA funding for ongoing NEMI development and maintenance will be leveraged so that CDI funding can be devoted to development of the protocol compendium extension.