NSDI Cooperative Agreements Program
Category 2: Best Practices in Geospatial Service Oriented Architecture (SOA)

Interim Report

Date: October 2, 2008

Agreement Number: 08HQAG0035

Project title: Documenting Best Practices in Geospatial SOA through the Development of a Wetlands Permitting Solution

Organization:
Image Matters LLC
201 Loudoun Street, SW
Leesburg, VA 20175
www.imagemattersllc.com

Project Leader:
John Davidson
V. 703.669.5510
F. 703.669.5515
johnd@imagemattersllc.com

Collaborating Organizations:
Environmental Protection Agency
Tod Dabolt, Director, Project Management Office, Office of Water
Voice: 202 564-1450
FAX: 202 564-0500
Dabolt.Thomas@epamail.epa.gov
http://www.epa.gov/ow/

U.S. Fish & Wildlife Service
Bill Wilen, Senior Biologist, National Wetlands Inventory
Voice: (703) 358-2161
FAX: (703) 358-2232
Bill_Wilen@fws.gov
http://www.fws.gov/nwi/

Executive Summary

The project is designed to fulfill two primary goals: (1) develop and integrate Web services to enhance the capabilities of the DARTER, a SOA-based tool that fulfills the EPA and ACE's CWA Section 404 missions, in a manner that supports the FWS mission of NWI data provision, and (2) produce guidance materials that document the SOA process for business process modeling, use case and requirements elucidation, service design, development, implementation, integration, testing and operational deployment of the Wetlands Screening Application. To date, services that perform "intersection" and "proximity" geo-analytical processes through an OGC WPS interface have been built and tested, as well as a client, the "GeoAnalysis Tool" application, which accesses and consumes USGS National Hydrographic Dataset (NHD) and USFWS National Wetland Inventory (NWI) features via OGC WFS. Preliminary documents have been produced that describe the SOA-based service development process. Our approach and that of our two co-awardees was presented at the GIScience 2008 conference in Park City, UT on September 23rd.

Project Narrative

After the project KO, Image Matters established a wiki-type collaborative environment (http://cap.imagemattersllc.com/confluence/) for FGDC and the other two grant recipients, IU and CubeWerx. This "Confluence" site has been used to develop and archive project artifacts including "common terms" used in SOA development, guidelines for the development approach and SOA modeling environment, and examples of documentation from each of the three projects. The content and structure of the "Confluence" site, and possibly the site itself, will be made available to federal agencies and others who seek guidance on how to build a SOA-based web service.

In addition to the Dashboard and project-specific pages, the site currently has the following primary pages, the first four of which are populated:

- Common Terms – a list of commonly-used terms, their definitions, and source information
- Use Cases – the set of use cases for each of the three projects
- Templates – documents that serve as templates for development efforts in other projects
- Best Practices – guidance materials for federal agencies looking to develop SOA-based web services
- Lessons Learned – shared experience for agencies looking to develop SOA-based web services
- Open Issues – areas that require further investigation and resolution, to be addressed by the three projects

With respect to its proposed project, Image Matters has followed the development process (discussed below) in building an application which accesses and consumes US Geological Survey (USGS) National Hydrographic Dataset (NHD) and US Fish & Wildlife Service (FWS) National Wetland Inventory (NWI) features via OGC WFS, and performs "intersection" and "proximity" geo-analytical processes through an OGC WPS interface. We are currently working on the following aspects of the application, given the generic name of "GeoAnalysis Tool":

- Performance – the access to the USGS NHD feature-level data appears to the rate-limiting step in the process. We are looking for alternative sources of the same information and ways to speed up access to the existing server.
- Ancillary Data Additions – we plan to integrate additional ancillary data to serve as context for the EPA Wetland Analysts who would use this tool. Currently, we have selected a pilot study region in the Choptank watershed area of Maryland for which LiDAR and additional wetland delineation layers exist. We will be setting up a temporary WMS for the LiDAR data to allow its integration into the map.
- Minor Enhancements – following review of the application, we anticipate making minor enhancement based on user feedback

The project has run into two roadblocks with respect to its original goals. The first of these impediments centers on the integration of our geoprocessing services into the Data on Aquatic Resources Tracking for Effective Regulation (DARTER) application, which is targeted to be the next generation software framework for EPA's wetland analysts and permitting managers. DARTER development is behind the anticipated schedule, and is not ready for integration with our services within the original 10-month timeframe of our project. We still hope to be able to integrate our services within DARTER, and will adjust our schedule to the greatest extent possible to account for the unexpected delays. In the meantime, we were forced to build our own client software in order to demonstrate, receive feedback on, and further refine our services, as well as provide input to the final toolset requirements that will eventually be a part of the DARTER system.

The second roadblock involves finding a willing host for the FWS's NWI WFS. The planned deployment at an EPA server in Madison (WI) fell through due to internal EPA resource allocations. The alternative, for FWS to host the FWS from their own server, was untenable because of the perceived burden that it

would place on staff who are already stretched thin. Image Matters responded by standing up NWI data behind a WFS on one of their own servers, a temporary solution that allowed development to progress.

Fortunately, neither roadblock prevents our completion of software development and testing. Certainly there were valuable lessons learned through this process (and these will be documented), although not exactly the type we expected.

SOA Definitions and Approach

Image Matters generated an initial list of 33 commonly-used terms (presented below) and their definitions and posted these to the "Confluence" project collaboration site for consideration by FGDC and the two other CAP2 award recipients. The terms and definitions were taken from authoritative sources, and the references to those sources are included in the listings. We will continue to add to the list and refine the individual definitions throughout the duration of the project.

Table 1. Common Terms defined to date

Atom Publishing Protocol	literal	service
Atom Syndication Format	literal (XML encoding)	service chain
Blog	map	service instance
bounding box	Microformats	service metadata
capabilities XML	operation	Service Oriented Architecture (SOA)
client	output	
GEO Microformat	parameter	SOAP
geographic information	process	version
GeoRSS	request	Web Service Architecture (WSA)
interface	response	
interoperability	REST	workflow
	RSS	WSDL

Requirements and Process Definition

Image Matters is following the general software development life cycle pattern agreed upon by the three awardees: model process and elucidate requirements, design and develop, implement and test, deploy and monitor. The first step has been broken down further into the following components:

1) Document Business Process
2) Create Concept of Operations
3) Develop Detailed Use Cases
4) Generate Technical Requirements

Our requirements gathering phase started, albeit informally, during the proposal formulation stage. At that time we were made aware of a specific geo-analytical processing need that would best be addressed with a SOA-based web service. After researching the definition of a "jurisdictional wetland" and confirming the correct inputs and outputs, we proposed our solution which was met favorably by the EPA and FWS.

Upon receiving funding, we requested, received, and examined the multi-agency business process needed to make Jurisdictional Determinations (JD), as prepared for the DARTER development. EPA's

"Screening" of US Army Corp of Engineers (COE) preliminary decision regarding the permit was step 13 of a 24-step process. So while the business process placed the Screening Step, in which all EPA geospatial analysis took place, in the context of the larger business process, the process models were at too high of a level to provide much information about the specifics of geoprocessing required by EPA's Analysts for Jurisdictional Determination (JD). It was clear the specific step-by-step process involving geospatial data and geo-analytical processing required by the EPA analysts needed to be captured. We determined that this level of detail was best documented in a use case format.

As part of our first meeting with all involved parties, including program managers, data specialists, and wetland analysts from EPA, FWS, the US Army Corp of Engineers (COE) and consultants leading the DARTER development effort, we spent time gathering specific user requirements. We validated our proposed geoprocesses, and received input on necessary user controls, formats for presenting results, and desired contextual layers. The information gathered at that meeting provided us sufficient information to develop detailed use cases for software interactions.

Currently, the constituent user base consists solely of EPA Wetland Analysts. However, at our requirements meeting, the COE expressed an interest in enabling their wetland analysts to access the WPS-based geoprocessing services, which are a good fit with the SOA-based design of the ORM-II (OMBIL Regulatory Module). [Note that the EPA's DARTER system is a part of the larger ORM-II system, and the two software frameworks exchange information as a part of the multi-agency Jurisdictional Determination workflow.] In addition, it is quite possible that States, Tribes, an even municipalities would find the geoprocessing services useful for landscape-level or site-specific assessment of wetland connectivity to water features. Given that the WPS-based services were designed to be generic, custom applications could be developed that utilized inputs other than NHD and NWI, e.g., high-resolution wetland delineations generated from field GPS work and stood up via WFS. Moreover, assuming adequate access and WFS-based data inputs, users faced with any intersection and/or proximity geoprocessing requirement involving web services could utilize these services.