GEO-PROCESSING IN CYBERINFRASTRUCTURE:
MAKING THE WEB AN EASY TO USE GEOSPATIAL COMPUTATIONAL PLATFORM

George Percivall[a,*], Lionel Menard[b], Lan-Kun Chung[c], Stefano Nativi[d] and Jay Pearlman[e]

[a] Open Geospatial Consortium, Crofton, MD 21114, USA – gpercivall@opengeospatial.org
[b] MINES ParisTech, Sophia Antipolis Cedex, FR – lionel.menard@mines-paristech.fr
[c] GIS Research Center, Feng Chia University2 Taichung City 407 Taiwan – peter@gis.tw
[e] IMAA-CNR, Potenza, Italy – nativi@cnr.it
[e] IEEE, Seattle WA 98126, USA – jay.pearlman@ieee.org

Abstract — **Access to data on the web has become routine based upon open standards from IETF and W3C. Access to explicitly geospatial data is routinely done using data access standards from the OGC. Geoprocessing services on the web are now being developed. Processing of data must be done to apply or fuse the data to meet specific applications. Standards and implementations for processing of data on the web are just now becoming established. For geospatial data, the OGC has defined the Web Processing Service (WPS) interface standard. Now is a critical time to bring convergence to WPS profiles that make the web an easy to use geospatial computational service. Access to network accessible processing services is bringing geo-processing to the cyberinfrastructure.**

Keywords – **geoprocessing, interoperabilty, standards, web services, data fusion, algorithms, GEOSS, biodiversity, disaster management, energy, socio-economic**

1 INTRODUCTION

Access to information on the web has become routine using open standards from World Wide Web Consortium (W3C). Access to explicitly geospatial data is routinely done using data access standards from the Open Geospatial Consoritum (OGC), e.g. Web Map Service, Web Feature Service, Web Coverage Service. As the maturity of web access to geospatial data has been achieved with open standards, access to processing services on the web has been developing.

Previously, most of a scientist's time was spent gathering and preparing data before research could begin. Much of the data on the web is not directly useful to the end user's needs. Processing of the data must be done to apply or fuse the data to meet specific applications. Standards and implementations for processing of data on the web are just now becoming established for cyberinfrastructure. For geospatial data, the OGC has defined the Web Processing Service (WPS) standard.

The Group on Earth Observations (GEO) is using WPS in the development of the Global Earth Observing System of Systems (GEOSS). Results in GEOSS show the value of using WPS for geoprocessing. Further work in standardizing profiles of WPS will increase the interoperability between independently developed software and result in better geospatial information to inform research and decision-making.

2 WEB SERVICES FOR GEOSPATIAL INFORMATION

Today, when people think of access to information they expect to use the Web. Earth Observation (EO) information systems have made major advances to meet these expectations. Access to maps is almost completely done by web protocols. Access to EO data is advancing to

become dominated by the Web, although file-based order and delivery is still the norm for some. Web based access to EO data allows for the next step in meeting user expectations. Web based access allows for the development of geoprocessing of EO data on the web to meet user expectations of a robust cyberinfrastructure.

GEOSS provides an excellent example of the progress of web services for geospatial information. GEOSS aims to provide comprehensive, coordinated and sustained observations of the Earth. The GEOSS 10 Year Plan defined a bold vision that is now being realized by the Group on Earth Observations through an architecture based on interoperability arrangements.

The GEOSS 10 Year Plan clearly identifies the role of standards in the development of the GEOSS information system architecture:

> "The success of GEOSS will depend on data and information providers accepting and implementing a set of interoperability arrangements, including technical specifications for collecting, processing, storing, and disseminating shared data, metadata and products. GEOSS interoperability will be based on non-proprietary standards, with preference given to formal international standards." (GEO 2005)

The GEOSS Architecture Implementation Pilot (AIP) has defined and tested an architecture that provides access and processing based on open standards (Figure 1). Clients can directly obtain data from the Access Services or request mediation services. A main element of the mediation is processing of the geospatial data into information suitable to the needs of the client users. GEOSS has deployed several servers using the OGC WPS to produce information for users in several GEOSS Societal Benefit Areas: Energy, Biodiversity, Disaster Management and Socio-economic.
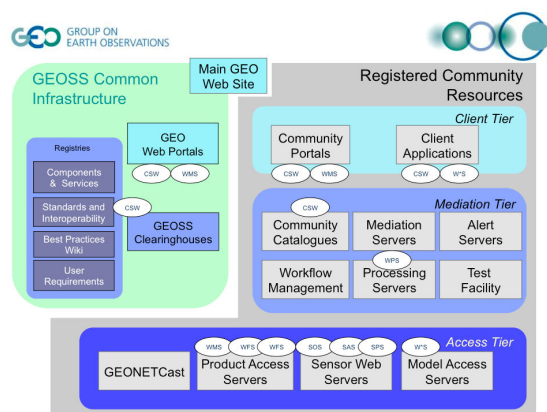

Figure 1. GEOSS Engineering Components provide access to services through standard interfaces.

3 THE NEED FOR GEOPROCESSING

In many instances geospatial information is not directly fit for the purpose of users. Geospatial measurements are stored in archives mainly from the perspective of how they were collected. The geophysical parameters and the spatial structure of the data reflect how they were defined for a given application. Instrument developers and archive managers strive to make the data as widely applicable as possible, but cannot possibly anticipate all users' needs. To be most effective, parameters or structure of the data may need to be changed or fusion of the data with other data may be needed.

Data fusion is *"the act or process of combining or associating data or information regarding one or more entities considered in an explicit or implicit knowledge framework to improve one's capability (or provide a new capability) for detection, identification, or characterization of that entity"* (Percivall 2010). Fusion techniques depend upon the level of fusion e.g., observation, feature, or decision fusion. Hundreds if not thousands of algorithms exist for the processing and fusion of geospatial data. For example, feature fusion includes techniques for identifying, aggregating, relating, parsing, and organizing and includes feature processing such as generalization, conflation, feature extraction, and change detection. Similarly, image processing has many different classes of algorithms: convolution, feature extraction, frequency filters, geometric operations, spatial filters and orthorectification. Processing to provide thematic, statistical, exploratory, spatial/topological and other forms of analysis are needed for socio-cultural analysis of geospatial information. (Kiehle 2010).

Geoprocessing Workflow is an automation of a spatial process/model, in whole or part, during which information is passed from one distributed geoprocessing service to another according to a set of procedural rules using standardized interfaces (Schäffer, 2009). Combining processing and rules services within a service-oriented architecture with workflow is an enabling technology for fusion. The rules can be inspected and compared and subsequently executed on a variety of workflow processing services. This has been demonstrated for conflation and topology quality assessment (Werling 2008) and for Earth Observation data and processing (Falke 2008)

Ready access to processing services is needed for their individual use or when used in a processing workflow. The need is clear for uniform interfaces to enable ready processing on the web.

> "Developing a system that utilizes existing or developmental data fusion technology requires a standard method for specifying data fusion processing and control functions, interfaces, and associated data bases. The lack of common engineering standards for data fusion systems has been a major impediment to integration and re-use of available technology." (Steinberg 1999)

4 OGC WEB PROCESSING SERVICE

The OGC WPS standard defines a framework in which organizations are now making geoprocessing services more readily available over the web. Initial profiles of WPS have been developed for coordinate transformation, image processing, feature analysis and statistical analysis. Proprietary and open source projects have implemented WPS profiles. Now is a critical time to bring convergence to a minimal set of WPS profiles that can make the web an easy to use geospatial computational service. Access to

network accessible processing services is bringing geo-processing to the cyberinfrastructure.

The OGC Web Processing Service Interface Standard (OGC 2007) provides rules for standardizing communication of inputs and outputs (requests and responses) for geospatial processing services, such as polygon overlay. The standard also defines how a client can request the execution of a process, and how the output from the process is handled. It defines an interface that facilitates the publishing of geospatial processes and clients' discovery of and binding to those processes (Figure 2). The data required by the WPS can be delivered across a network or they can be available at the server. The WPS defines three operations that are made accessible to a client (Table 1).
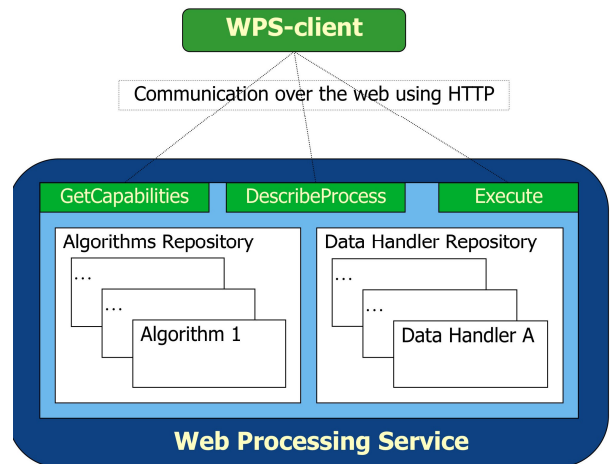

Figure 2. Context of WPS showing operations and implied repositories for algorithms and data handlers

Table 1. OGC Web Processing Service operations

Operation	Description
GetCapabilities	This operation allows a client to request and receive back service metadata (or Capabilities) documents that describe the abilities of the specific server implementation. The GetCapabilities operation provides the names and general descriptions of each of the processes offered by a WPS instance. This operation also supports negotiation of the specification version being used for client-server interactions
DescribeProcess	This operation allows a client to request and receive back detailed information about the processes that can be run on the service instance, including the inputs required, their allowable formats, and the outputs that can be produced.
Execute	This operation allows a client to run a specified process implemented by the WPS, using provided input parameter values and returning the outputs produced

OGC approved the WPS standard several years ago. At the time of approval it was well known that the WPS Standard would serve as a framework for the deployment of numerous (100's) of geoprocessing algorithms in distributed information networks. Multiple implementations of WPS are now available (OGC 2011), (FOSS4G 2010), (WebMGS 2010), (Wikipedia 2011).

The OGC is now undertaking deployments of WPS coordinated through WPS Profiles. A current challenge is to have consensus development of WPS profiles that encourages the highest level of interoperability. Profiles need to be general enough to address multiple implementations of a class of algorithms, while being

specific enough to provide interoperability between independently developed clients and servers

Recent work on feature and statistical analysis processing with WPS identified hierarchies and classifications such as "Topology Analysis" and "Statistical Analysis" (Kiehle 2010). Conflation rules also seem to be a general class that should be profiled in WPS (Brennen 2008).

Continued implementation of WPS should be pursued to inform the judicious choice of WPS Profile development. It is critical that WPS profiles begin to emerge for popular classes and grow in consensus in order to avoid the creation of a separate profile for every algorithm implementation.

5   APPLICATIONS OF WPS IN GEOSS

The GEOSS Architecture Implementation Pilot (Figure 1) includes application of WPS to several GEOSS Societal Benefit Areas (SBAs). These applications provide ready results for the SBA users and demonstrate that the WPS standard is effective in making geoprocessing readily available in a distributed web services environment based on open standards.

5.1   EuroGEOSS "eHabitat" use of WPS

The GEOSS AIP "eHabitat" scenario applies the WPS to analysis of the effects of Climate Change on Biodiversity. The eHabitat scenario applies the developments of the EuroGEOSS project funded by the European Commission. (Nativi and Santoro, 2011).

With the scenario of climate change impact on protected areas in mind, a Web based decision-making tool for assessing potential environmental changes due to anthropogenic activities, including climate change was developed. In particular, the development with eHabitat of a modeling web service allowing end-users to assess the likelihood of finding equivalent habitats, considering a virtual infinity of possible ingredients defining these habitats, provides decision makers with useful information on the irreplaceability of the modeled habitats.

The ability of eHabitat, as an open source Web Processing Service (WPS), to be a component of any processing chain, combined with the possibility to use any thematic map to define a model of the habitat requires the architecture to be interoperable from an inter-disciplinary point of view. The WPS must also be able to cope with multi-scale heterogeneous data sources.
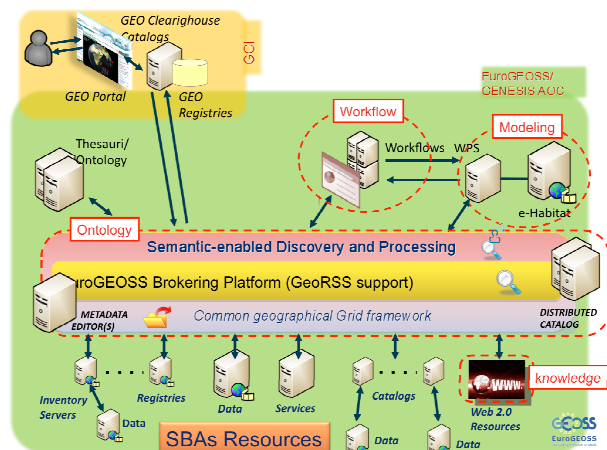

Figure 3.   EuroGEOSS WPS to model of biodiversity habitats.

5.2   MINES ParisTech use of WPS for Energy

The GEOSS AIP Energy scenario provides spatial information on the life cycle environmental impacts of the production of photovoltaic electricity. Renewable energies do contribute to the reduction of GHG compared to fossil fuels, but carbon is created through manufacture, maintenance and other facets of operations. The scenario focuses on the assessment of such impacts for photovoltaic systems by a proper exploitation of data available within GEOSS. (Menard, 2011)

Life Cycle Assessment information of Photovoltaic Systems must be understood and quantified to report on the environmental impact of PV systems regarding the complete supply chain of each system taking into account geo-localized solar radiation parameters. OGC WMS and WPS were selected as standardized means to access and process such information. Based on Life Cycle Inventories of PV systems (the first step in a life cycle assessment calculation) that have been provided by ecoinvent, MINES ParisTech developed Web Services that encapsulate the environmental assessment computation process. These Web Services have been deployed on the Energy Community Portal (www.webservice-energy.org) and deployed in GEOSS AIP-3. (See Figure 4)
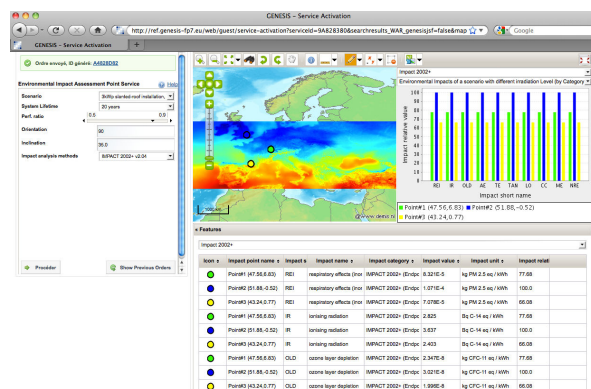

Figure 4.   Results of the AIP-3 "Point Based Scenario" invoking remote WPS through a Web based graphical client. Geolocalized enviromemental impacts parameters of PV systems are displayed using matrix and bar graph.

Figure 5.

5.3   GIS FCU use of WPS for Emergencies

The GEOSS AIP Disaster Management Working Group implemented flooding disaster management service as a specific instance of a Disaster Management Reference Scenario. The GIS Center of Feng Chia University (GIS FCU) developed a WPS for shortest path calculation for dispatch of emergency vehicles. (Chung and Cauchy, 2010)

This scenario includes integration and utilization of GEOSS standard components and services to supply the near-real-time dispatching of emergency vehicles during the response phase. The scenario is applied to flooding disasters caused by tropical storms, hurricanes, cyclones, or typhoons in particular, but can be easily re-cast to cover other disaster types such as earthquakes, wildfires, landslides, volcanoes, and tornadoes.

WPS is used to run an algorithm to find the optimal solution under the constraints of a flooded area and a near-real-time traffic report, resulting in the final shortest path being delivered to an integrated client.

The scenario illustrates how services in GEOSS can meet the requirement of producing a near-real-time dispatching map based on the comprehension judgment fused by various sensors services in order to help the dispatching

decision maker to make a rational decision for dispatching an emergency vehicle from a specific origin point to a particular destination.
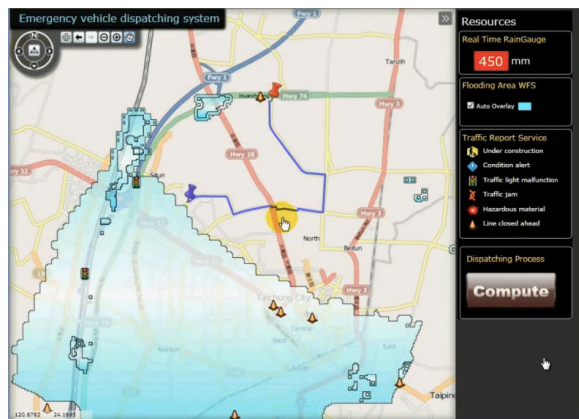

Figure 6.   Results of GIS FCU WPS dispatch around flood

5.4 CIESIN's use of WPS for Population Analysis

As part of the GEOSS AIP-2, CIESIN set up a GIS analysis of raster population data (persons, 2005) as a WPS. The WPS accepts simple feature (polygon) GML (i.e., data encoded using the OGC Geography Markup Language Encoding Standard) from a WFS (i.e. a server implementing the OGC Web Feature Service Interface Standard) or as a file upload and returns the same features with associated population statistics (parametric statistics, population totals, and data quality indicators). The WPS acts a proxy to an ArcGIS Server Geoprocessing task. The Population Statistics WPS is intended to be used for analysis of GIS data or as a single component contributing to a more complex workflow.

A demonstration of the Population Statistics WPS is available as a video. See "Geoprocessing Services" video: http://www.ogcnetwork.net/pub/ogcnetwork/GEOSS/AIP2/index.html.

6 CONCLUSION

Meeting and exceeding user expectations for ready access and processing of Earth Observation data is well underway through the use of open standards. The open standards of OGC and other organizations have been deployed and used to provide robust, full-functioning cyberinfrastructure based on web services. The OGC Web Processing Service is the next major element to meeting the users needs, allowing for interoperable access to hundreds of algorithms for geospatial information.

The use of cyberinfrastructure standards for data access and geoprocessing is vital to achieving the vision of GEOSS. The pioneers in GEO are making vast amounts of Earth Observation data readily applicable to solving difficult problems in the Societal Benefit Areas. Contributions from hundreds of organizations - providing reusable components and services - has led to substantial architectural achievements, cross-cutting all Societal Benefit Areas.

GEOSS is making Earth observations more readily accessible and useful to more people. Progress on the bold vision of GEOSS defined 5 years ago is making a difference in our world.   The ceaseless collection, processing, and application of Earth Observations is vital to addressing humanity's critical decisions.

REFERENCES

(Brennen 2008) "OWS-5 Conflation" Engineering Report, OGC Document 07-160r1, Pete Brennen, September 12, 2008.

(Chung and Cauchy, 2010) "Disaster Management Scenario" Engineering Report, GEO Architecture Implementation Pilot, Phase 3, Lan-Kun (Peter) Chung and Arnaud Cauchy, Version 1.4, December 24, 2010.

(Falke 2008) "OWS-5 Earth Observation Web Processing Services (WPS)" Engineering Report, OGC Document 08-058r1, Stefan Falke, September 12, 2008.

(FOSS4G 2010)   Presentations about WPS at FOSS4G conference, September 2010, Open Source Geospatial Foundation (OSGeo), http://2010.foss4g.org/presentations_gen_sel.php

(GEO 2005) "Global Earth Observation Systems of Systems – GEOSS – 10-Year Implementation Plan", GEO 1000 / ESA BR-240, February 2005, Published by ESA, The Netherlands

(Kiehle 2010) "OWS-7 Web Processing Service Profiling" Engineering Report, OGC Document 10-059r2, Christian Kiehle and Theodor Foerster, August 18, 2010

(Menard, 2011) "Energy Scenario" Engineering Report, GEO Architecture Implementation Pilot, Phase 3, Lionel Menard, Version Final, January 6, 2011.

(Nativi and Santoro, 2011). "eHabitat - Climate Change and Biodiversity WG Use Scenario" Engineering Report, GEO Architecture Implementation Pilot, Phase 3, S. Nativi and M. Santoro, Version 2.0, January 7, 2011,

(OGC 2007) "OpenGIS® Web Processing Service," Open Geospatial Consortium Document 05-007r7, Perter Schut Editor, Version 1.0.0, June 8, 2007.

(OGC 2011) Implementations by Specification – WPS 1.0, OGC Web page accessed 28 January 2011. http://www.opengeospatial.org/resource/products/byspec

(Percivall 2010) "OGC Fusion Standards Study, Phase 2" Engineering Report, OGC Document 10-184, George Percivall, December 13, 2010.

(Schäffer, 2009). "OWS-6 Geoprocessing Workflow Architecture" Engineering Report, Open Geospatial Consortium document 09-053r5, Bastian Schäffer, October 9, 2009.

(Steinberg 1999) "Revisions to the JDL Data Fusion Model", Steinberg, A. N., Bowman, C. L., & White, F. E.. ERIM International, Inc., 1999

(WebMGS 2010).   Presentations about WPS at WebMGS 2010: 1st International Workshop on Pervasive Web Mapping, Geoprocessing and Service, Como, Italy, August 26 to 27, 2010 http://webmgs2010.como.polimi.it/

(Werling 2008) "OWS-5 GeoProcessing Workflow Architecture" Engineering Report, OGC Document 07-138r1, Michael Werling, September 12, 2008-09

(Wikipedia 2011) Wikipedia article about WPS. Web page accessed 28 January 2011 http://en.wikipedia.org/wiki/Web_Processing_Service

(Yetman, 2009) "Workflow and Processing" Engineering Report, GEO Architecture Implementation Pilot, Phase 2, Greg Yetman, CIESIN, Version 1.0, September 9, 2009.