GEO GRID Platform for Integrated Earth Sensing

Sarawut Ninsawat
School of Engineering and Technology
Asian Institute of Technology, (AIT), Thailand
sarawutn@ait.ac.th

Yoshio Tanaka and Hirokazu Yamamoto
National Institute of Advanced Industrial Science and Technology (AIST), Japan
{yoshio.tanaka,hirokazu.yamamoto}@aist.go.jp

Sameer Tilak and Peter Arzberger
University of California, San Diego, USA
{stilak,parzberg}@ucsd.edu

***Abstract*— An Integrated Earth Sensing demands seamless access to large-scale geospatial data gathered by remote sensing and in-situ sensor network deployments. To that end, the Global Earth Observation (GEO) Grid provides a framework for sharing a wide array of scientific information such as satellite data, geologic maps, and in-situ sensing data in a standards-compliant manner. In this paper, we describe three applications of GEO Grid in the domain of environmental sensing, disaster management, and largescale geo-science data provisioning. Our approach uses web service technologies to promote interoperability and wider dissemination of technology and data.**

Keywords*— *Integrated Earth Sening, sensor networks, remote sensing, Open Geospatial Consortium (OGC) Web services

I. INTRODUCTION

Remote (Satellite and airborne based), and in-situ sensing are three prominent ways to generate unprecedented synoptic coverage of the Earth (ref. Figure 1). Satellite Earth observation sensors provide unique measurements of geophysical and biospheric variables globally and repetitively In situ measurements make it possible to supplement and ground-truth validate satellite sensor observations and there will always be parameters that are inaccessible from space. Integrated Earth Sensing to monitor remote environments, hazards and disasters, and natural resources using remote and in situ measurements. The knowledge gleaned from Integrated Earth Sensing has the potential to empower managers and decision makers to act on critical climate, sustainable development, natural resource, and environmental issues [1]. To that end, The Global Earth Observation (GEO) Grid provides a platform for sharing earth observation data, storage, and computational powers of high performance computing and is accessible as a set of services using the grid technology. It is a robust IT infrastructure, which federates distributed and heterogeneous Earth observation data [6]. In this paper, we describe how the GEO Grid framework has been used to enable a wide range of applications from environmental sensing/modeling to disaster management.

II. GEO GRID ARCHITECTURE AND APPLICATIONS

GEO Grid is based on four layers, hardware, virtual storage, application and data services, and user interface. Each layer has the potential for extendibility (ref. Figure 1).

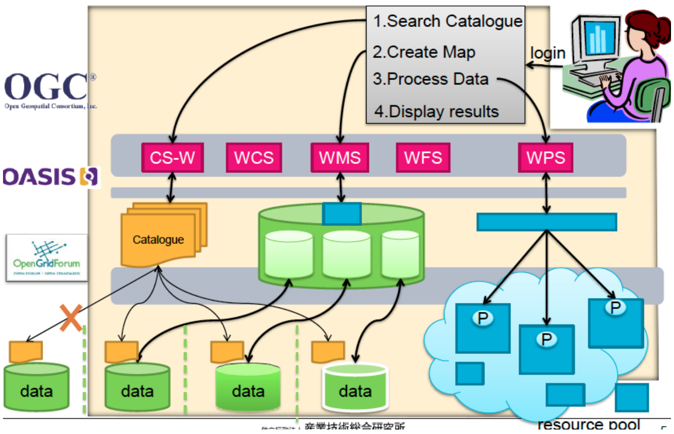

Figure 1: GEO GRID Architecture

The Global Earth Observation (GEO) Grid [6] is aiming at providing integrated services using a wide array of Earth scientific information such as satellite data, geologic maps and other digital geographically referenced data, and assembles them into easy-to-use formats for potential stakeholders from several fields such as environmental conservation, resource exploration, hazard evaluation, and risk management.

The "GEO Grid" project, since 2005, is primarily aiming at providing an E-Science infrastructure for worldwide Earth Sciences community. In the community there are wide varieties of existing data sets including satellite imagery, geological data, and ground-sensed data that each data owner insists own licensing policy. Also, there are so many of related projects that will be configured as virtual organizations (VOs) enabled by Grid technology. The GEO Grid is designed to integrate all the relevant data virtually, again enabled by Grid technology, and is accessible as a set of services. In this paper we describe following three key applications. related to the earth observation using the grid technology, which is developed for sharing data, storage, and computational powers of high performance computing, and is accessible as a set of services.

- Satellite data application : Application of Satellite-Field data Integrator (SFI) for Satellite and in-situ sensor data integration
- Hazard information : QuiQuake (Quick Estimation System for Earthquake Maps Triggered by Observation Records)

  http://qq.ghz.geogrid.org/QuakeMap/index.en.html
- Large-scale Geoscience data sharing: Geological maps, Active fault data for providing large-scale geospatial data (raw and maps).

We now describe our experience in developing and deploying these applications using the GEO Grid framework.

A. Satellite-Field data Integrator (SFI): Satellite and in-situ sensor data integration

Satellite-Field data Integrator (SFI) is a platform for integrating remote sensing data with in-sit sensing data and exposing that to end users in a standards-compliant manner. An integrated earth observing system therefore needs seamless integration of both the remote sensing and in-situ sensing data. We now enumerate the pros and cons of these two sensing methodologies.

Both the remote sensing and in-situ sensing paradigms have pros and cons. The main benefits of satellite/ remote sensing are (1) Remote sensing image of the earth observation is an efficient method to monitor the environment. In particular, it allows wide spatial coverage at cheaper cost (2) The capability to collect multi-spectral imagery of the same area at different periods of time allows us to detect the change of environment over time. Regional coverage and broadly spectral resolution, (3) Continuous acquisition of data, and (4) Archive of historical data. On the other hand, following are the limitations of satellite/ remote sensing are (1) Interference of atmospheric gaseous and particles Absorbing (H20, O3 etc.) and Scattering (mainly by aerosol particles such as dust, ash and smoke), and (2) By definition direct sample of the phenomenon is not possible. On the other hand, in-situ sensing has the following advantages: (1) Direct or similar sample of the phenomenon (2) Real-time or Near Real-time observation (3) High temporal resolution. Unfortunately, the main limitation of in-situ sensing the fact that it is expensive for wide area observation.

As mentioned before, the reliability of remote sensing data itself affected by physical constraints such as the absorption and scattering of atmospheric molecules and aerosols. The accuracy of quantitative modeling of observed spectral to phenomenon parameters is critical. The errors caused by remote sensing data are mostly disregarded in many studies as the major concern adequate accuracy Thus, the validation process is an essential task for geo scientist to assess the quality of satellite products prior to use it for further analysis. Not only assessing the quality of satellite product is possible by the validation activities with ground measurements, but also the gathering of information that is scarcely obtained from direct satellite observation is feasible by the relationship evaluation process.

Various techniques for validation and correlation evaluation tasks have been carried among differ group of scientist. However, most of common tasks are the same especially in the part of data management. In particular, analysing and processing of both satellite data and ground measurement involves a significant effort in terms of searching, acquiring, cleansing, and processing volumes of data. The complexities involved in this workflow demand sophisticated user skills and resource-rich facilities. Unfortunately this limits the availability of this important data to the broader user community. To that end, in this paper we describe our approach to address the big challenge of developing an e-Science infrastructure to simplify the usage of remote sensing data to users. Our approach uses web service technologies and interoperability standards. Since interoperability is increasingly becoming a focus point for organizations that distribute and share data over the Internet, our approach helps wider dissemination of data. We now describe the details of The Satellite-Field data Integrator (SFI) framework that runs on top of the GEO Grid platform.

The Satellite-Field data Integrator (SFI) framework is designed to reduce the onerous tasks of data gathering, manipulating, and processing. It supports heterogeneous data formats in both remote sensing and sensor observation data. It is designed to handle the increasing number of datasets currently available in a scalable manner and it also offers a robust, on-demand processing service. The key components of SFI for providing interoperable, search and on-demand data and processing capabilities access as web services is conducted by various open standards of Open Geospatial Consortium (OGC) Web Service specifications [9][10][11]. As shown in Table I, the lists of available OWS specifications are used to fulfill functional requirement.

TABLE I. LIST OF OWS IMPLEMENTED IN SFI FRAMEWORK

OWS Specification	**Functionalities**
Web Mapping Service (WMS) Web Coverage Service (WCS)	Accessing to Satellite Observation and preview results map data
Sensor Observation Service (SOS)	Accessing to *in-situ* measurement
Web Processing Service (WPS)	On-demand processing/calculating capabilities

In SFI framework, the in-situ measurements are used as the reference data. The SOS web service is used to enable scalability for wide area sensor deployment and interoperability despite the presence of heterogeneous sensing data, manufacturer and network infrastructure. The GetObservation request of SOS is a main interface that allows users to access observation data. The observation data can be screened with several filter expression for example spatial location, station name, observed property and temporal period.

The WMS and WCS are used as the core web service that allow user to access satellite product. A user overcomes the complex tasks for satellite image processing by issue a standard request via TCP/IP protocol. The time-series observations at a particular location of satellite image are accessed by the GetFeatureInfo interface of WMS. The returned text-based data is used to validate with the in-situ observation. In contrast with the returned result as text-based format by GetFeatureInfo of WMS, the GetCoverage of WCS is used for standard data sharing to enable access to raw values or properties of coverage data and return it as coverage or raster image format. The results of request can then be processed for further analysis in any GIS software on user machine or other client services.

The on-demand processing capabilities are afforded by WPS web services. The preprogram processes including validation, evaluation of relationship, statistical analysis and processing for desired map were prepared in WPS system. The data source from WMS, SOS and WCS will be requested as the input data to complete the process as shown in Fig. 3. The final result will be shown in the web application as the preview map and users can download it for further analysis.

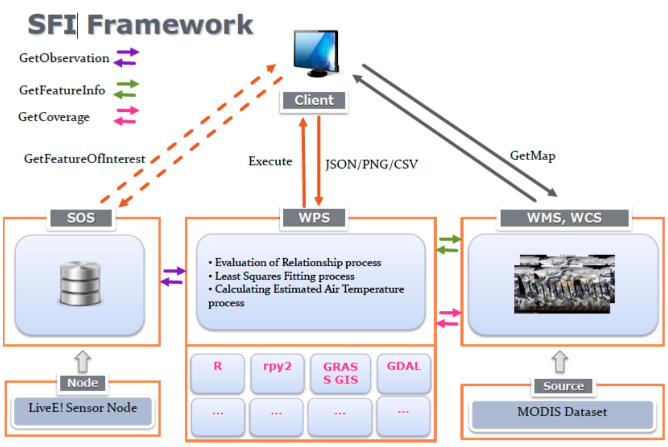

Fig. 1. SFI framework in GEO Grid Architecture GEO GRID Architecture

1) The Global Lake Ecological Observatory Network (GLEON)

GLEON is a grassroots network of limnologists, information technology experts, and engineers who have a common goal of building a scalable, persistent network of lake ecology observatories. Data from these observatories, including The Long Term Ecological Research [LTER] Network sites, will allow us to better understand key processes such as the effects of climate and land use change on lake function, the role of episodic events such as typhoons in resetting lake dynamics, and carbon cycling within lakes [5]. The North Temperate Lakes Long-Term Ecological Research (NTL-LTER) site is part of GLEON. The observatories will consist of instrumented platforms on lakes around the world capable of sensing key limnological variables and moving the data in near-real time to web-accessible databases.

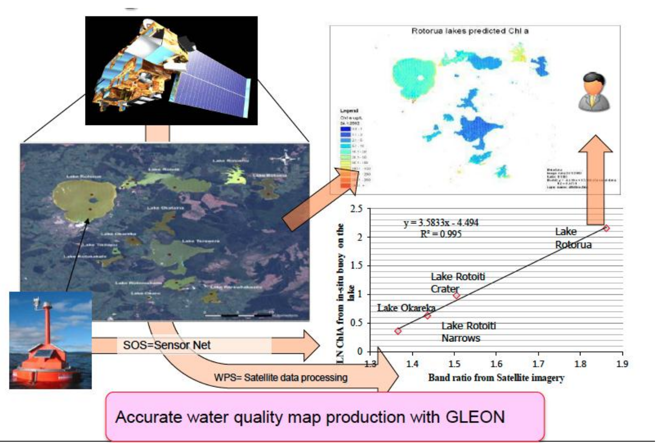

Fig. 2. Combining satellite and in-situ data

2) The Coral Reef Environmental Observatory Network (CREON)

CREON is a collaborating association of scientists and engineers from around the world striving to design and build marine sensor networks [4]. Sensor networks on land are becoming common as this technology allows us to view the environment in real time and in many dimensions. Extending this to the marine environment poses many challenges. However the benefits are enormous as we attempt to understand the stresses that are shaping the marine world. Coral reefs, in particular, are exhibiting signs of decay around the world as global warming; over fishing and pollution have an impact. The CREON group is presently deploying sensor networks in locations as diverse as the Moorea Coral Reef Long Term Ecological Research Site (MCR LTER: http://mcr.lternet.edu/), in Moorea, French Polynesia to the shores of Taiwan in the Kenting Coral Reef Group and the Great Barrier Reef. Using a variety of platforms and instruments the CREON group hopes to solve these challenges in a collaborative framework.

3) Real-world Experience of using SFI for lake monitoring by collaboration with GLEON sensor network

This application was developed as comprehensive web based application utilizing the SFI framework for lake monitoring. The satellite observation will be validated with in-situ measurement to improve uncertainly parameter measurement. In this application, the near surface water temperature is a key parameter for this development. The in-situ measurements from the observation site at lake Rotorua, New Zealand as a part of the Global Lake Ecological Observatory Network (GLEON) from July, 2007 to May, 2009 recorded every 15 minutes were used. Not only near surface water temperature measurements were used, but also other three water quality parameters (chlorophyll A, dissolve oxygen and water temperature at difference 13 depth level) and six weather measurements (wind direction, wind speed, air temperature, humidity, pressure and rain fall) were included. Total data is around 1.5 million records of observation are provided in the SOS which provided high efficiency capable of spatial, temporal and complex query. MODIS Level 2 Sea Surface Temperature product (MOD28/MYD28) at 1-km spatial resolution is the product of satellite observation. More than 700 raw data in HDF (Hierarchical Data Format) file format [12] were downloaded. The pre-processing including map re-projection, extract necessary data and file conversion were done to enable data in ready to used stage during the service operation. The web application is developed to provide the ease of use to users through common web browser with the rich interactive web mapping client to visualize and facilitate on-demand processing. User can select the study period then launch the process with supplementary special parameter. The return scatter plot between satellite observation and in-situ measurement will be shown with the related correlation parameters. The statistical analysis and fitting process are performed with the statistical analysis function of R software via the WPS process. Eventually, the sea surface temperature of MODIS product will be re-calculated with the input of slope and intercept of fitting equation. The estimated sea surface temperature process in WPS used the capabilities of Geographic Resources Analysis Support System (GRASS) Geographic Information System (GIS) [13] and GDAL library [14]. GRASS GIS is a free GIS software for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization. GDAL is an open-source library for reading and writing raster geospatial data formats. The final output will be available for user download after process finished. Current application offers data only at the lake Rotorua; however the scalability of developed framework is also possible for larger coverage. Since there are more than 20 observation sites as the part of GLEON, rather than the fitting process uses time-series observation at single site, we can apply the fitting process with multiple observation site at single date or apply the fitting equation to other lake that measurement equipment are not available.

a) Experience with integrating CREON data: We also imported subset of data from MCR-LTER CREON site into SFI. We were able to quickly import the data and expose it via SOS. Our exercise revealed that the developed architecture is quite flexible and it could seamlessly handle data from other grassroots networks such as CREON.

b) Estimating Air Temperature from MODIS LST and Sensor Network: We also applied SFI framework for estimating air temperature map from MODIS LST evaluated relationship with in-situ data collected over a distributed sensor network of ground sites in previous study [15]. The availability of high temporal measured air temperature by sensor network provided high-quality ground based data for estimating air temperature from satellite observation. The MODIS LST dataset as a data source from WMS server was evaluated relationship with near surface air temperature of Live E! Project [16] sensor network weather station from SOS server to calculate air temperature map with on-demand processing capability of developed WPS server requesting satellite data source from WCS server.

c) Lessons learned: (1) Satellite data although, large-in size, is homogenous in nature. Therefore integrating it is typically one-time effort. On the other hand, in-situ sensor data is very heterogeneous in nature, so even though small in size is quite heterogeneous and integrating its data is non-trivial and is typically an ongoing operation. (2) Use of standards is critical for integration, interoperability and wider data dissemination. (3) We were able to quickly develop and deploy SFI-based platform for estimating air temperature from satellite data and in-situ sensing data. This illustrates that the developed framework can easily support broad range of applications.

B. QuiQuake (Quick Estimation System for Earthquake Maps Triggered by Observation Records)

QuiQuake (Quick estimation system for earthquake maps triggered by observation records) [5] is a seismic event triggered quick estimation system for the generation of earthquake maps using earthquake observation records. It is one of the geological hazard related applications on GEO Grid and the first system which estimates and illustrates the extensive and detailed ground motion maps such as peak ground acceleration, peak ground velocity, and Japan Meteorological Agency-scale instrumental seismic intensity, right after an earthquake occurs. In this system, an amplification capability of ground motion and Vs30 average shear-wave velocity map, which is estimated from a 250-m grid cell map of the Japan Engineering Geomorphologic Classification Map (JEGM) [8], are also used for the spatial interpolation calculation to generate strong motion maps. The interpolation also takes into account the characteristics of soil conditions. The strong ground motion maps are automatically calculated and published through an Open Geospatial

Consortium (OGC) standard web service interface, right after the seismic observation data are released by the National Research Institute for Earth Science and Disaster Prevention (NIED). In this paper, we present the outline of QuiQuake, the data and algorithms used to generate strong ground motion maps.

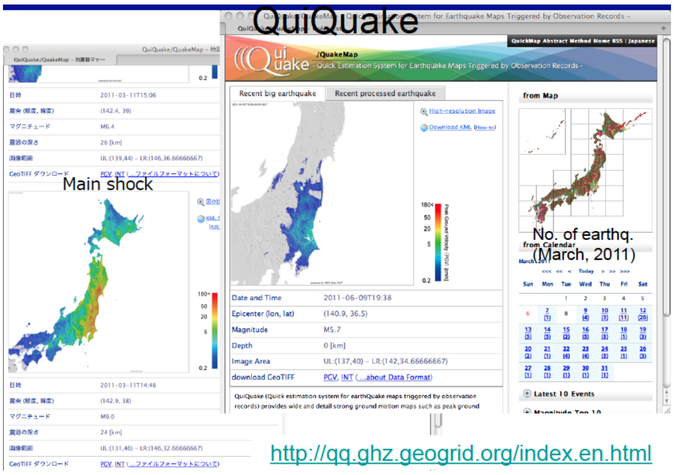

Fig. 3. QuiQuake

C. Earth science data services

GEO Grid provides wide variety of data services. Table II is a list of major data services provided by GEO Grid. As shown in Fig 5. shows how the ASTER (satellite) data can be quickly served in the form of map using OGC WMS service in case of a disaster.

TABLE II. A LIST OF MAJOR DATA SERVICES

Data	Description	Size	Format	Update frequency
ASTER	Satellite data observed by ASTER sensor	800 TB	Vector, Raster, GeoTIFF, etc.	Observed data is daily added to the database
Seamless geological maps	Seamless geological map of Japan with the scale of 1:200000 and 1:50000.	3.4 GB	Vector	Updated every year
Geological maps	Geological maps with the scale of 1:200000 and 1:5000000.	GB	Vector, Raster	Updated every year

IV RELATED WORK

The Flood Information Management System (FIMS) system combines real-time in situ data, internationally recognized hydraulic and hydrologic models, multi-resolution remotely sensed images and a customized Web interface [1]. With this system users are able to interactively visualize Ikonos, Radarsat, and Landsat image maps of their areas, visualize hydrometric station locations and actual flow data, then run models that show flood hazard, risk and extents based on current and predicted flows. Athos et al. [2] developed a system that integrated multi-spectral satellite images (Landsat 7 ETM+ and ASTER) and hydro-meteorological data from wireless sensors and automatic meteorological stations. These systems primarily focus on one specific domain such as and it is not clear how well it can support other applications. In addition, they do not focus on exposing data in a standards-compliant manner. In contrast, as described in this paper, GEO Grid framework has been used to support applications in multiple domains and both the sensor network and remote sensing data is exposed in a standards-complaint manner.

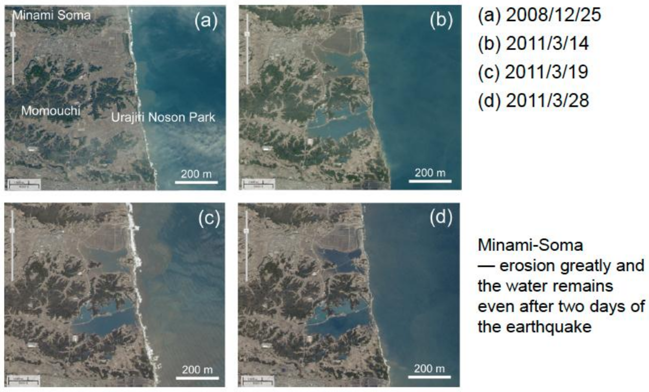

Fig. 4. ASTER Comparison Before and After Tsunami

V Future work

Even though the access to remote sensing data and ground measurement has been facilitated by the capabilities of services including, SOS, WMS and WCS, however the framework has been tested in few sites. We plant to deploy and test this framework to integrate network-wide data such as data gathered across all GLEON and CREON sites. We would also like to gather performance data and analyze the overhead of various Web services.

VI CONCLUSION

In this paper we described our experience with GEO Grid, an IT infrastructure, which federates distributed and heterogeneous Earth observation data. The proposed approach employs of Grid/cloud computing technology and exposes the acquired data using OGC standard to promote interoperability. Our experience shows that developing enabling real-world scientific applications that requires large-scale distributed and heterogeneous Earth observation data requires significant time and effort multi-disciplinary teams. Our experience also reveals that the use of standards promote extensibility, interoperability, and reusability. We also observed that integrating in-situ data, albeit small, compared to large-scale remote sensing data is non-trivial due to its heterogeneity. We plan to deploy SFI for network-wide studies. We also described QuiQuake, seismic event triggered quick estimation system for the generation of earthquake maps using earthquake observation records. Finally, we described how the GEO Grid framework was deployed to generate maps using satellite data to understand impact of disasters.

VII ACKNOWLEDGMENTS

This work is partially supported by NSF awards 1234983 and 1148458.

REFERRENCES

[1] G Wood, M.D.; Henderson, I.; Pultz TJ; Teillet, P.M.; Zakrevsky, J. G.; Crookshank, N.; Cranton, J.; Jeena, A., "Integration of remote and in situ data: prototype Flood Information Management System," *Geoscience and Remote Sensing Symposium, 2002*, vol.3, pp.1694,1696 June 2002.

[2] Agapiou, Athos, George Papadavid, and Diofantos G. Hadjimitsis. "Integration of wireless sensor network and remote sensing for monitoring and determining irrigation demand in Cyprus." *SPIE Europe Remote Sensing*. International Society for Optics and Photonics, 2009.

[3] Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst.* 15, 5 (Nov. 1993), 795-825. DOI= http://doi.acm.org/10.1145/161468.16147.

[4] CREON: The Coral Reef Environmental Observatory Network, http://www.coralreefeon.org/

[5] GLEON: The Global Lake Ecological Observatory Network, http://www.gleon.org/

[6] Sekiguchi, S., Tanaka, Y.;,Kojima, I., Yamamoto, N., Yokoyama, S., Tanimura, Y., Nakamura, R., Iwao, K., Tsuchida, S., "Design Principles and IT Overviews of the GEO Grid," *Systems Journal, IEEE* , vol.2, no.3, pp.374,389, Sept. 2008.

[7] Matsuoka, M., et.al., Web-based Quick Estimation System of Strong Ground Motion Maps using Japan Engineering Geomorphologic Classification Map and Observed Seismic Records, In Proceedings of 15th World Conference on Earthquake Engineering, pp. 4016-, 2012.

[8] Wakamatsu, K., et.al., Developing a 7.5-sec site-condition map for Japan based on geomorphologic classification., Earthquake Resistant Engineering Structures VIII, WIT Press, pp.101-112, 2011.

[9] Open Geospatial Consortium – Sensor Web Enablement Initiative http://www.opengeospatial.org/projects/groups/sensorweb

[10] Sensor Web Enablement Architecture, Category: Best Practice , Editors: Ingo Simonis, 2008.

[11] Mike Botts, George Percivall, Carl Reed, John Davidson OGC Sensor Web Enablement: Overview And High Level Architecture, OGC White Paper, 2007

[12] The HDF Group http://www.hdfgroup.org/

[13] Geographic Resources Analysis Support System GIS: http://grass.osgeo.org/

[14] Geospatial Data Abstraction Library: http://www.gdal.org/

[15] Sarawut Ninsawat, Hirokazu Yamamoto, Ryosuke Nakamura, Akihide Kamei, Seiichi X. Kato and Satoshi Tsuchida, "Development of OGC Framework for Estimating Air Temperature from MODIS LST and Sensor Network," 1st International Workshop on Pervasive Web Mapping, Geoprocessing and Services (WebMGS), 26-27 August 2010, Como, Italy

[16] Matsuura, S., Ishizuka, H., Ochiai, H., Doi, S., Ishida, S., Nakayama, M., Esaki, H. and Sunahara, H., 2007. Live E! Project: Establishment of Infrastructure Sharing Environmental Information. In Proceedings of the 2007 International Symposium on Applications and the Internet.