Major technical impediments to data sharing are not technical but political and cultural

Carl Reed, Chief Technology Officer at the Open Geospatial Consortium speaks to Defence IQs Asian Correspondent about the latest advancements and use of geospatial tools in Asia. He also outlines some observations on how the sector will evolve going forward.

Bryan Camoens: Could you please outline some of the latest advancements in the use of geospatial tools in Asia and how organization can incorporate them into their network?

Carl Reed: As a standards development organization, the Open Geospatial Consortium (OGC) does not develop tools. Instead, the members of the OGC are focused on the development of open geospatial encoding and interface standards known as encodings and interfaces that can be implemented in geospatial tools and applications. The goals are to enhance interoperability (information integration and fusion), reduce risk and lifecycle maintenance costs, as well as improve user experience.

The OGC now has 50 Member organizations in the Asia-Pacific region and the membership is growing. The interest and use of OGC standards in the region has increased significantly over the last several years. With the increased interest and activity, the OGC now holds one annual meeting in the Asia region. Last December, we had a Technical Committee (TC) meeting in Sydney, Australia. The TC is where all our members and non-members come together to discuss market trends, lessons learned on using OGC standards, and work on OGC standards. Our next TC meeting in June 2011 will be in Taichung, Taiwan. Current plans aim for OGC TC meetings to be hosted in Seoul, Korea in September 2012 and either China or India in 2013. Asia is becoming a very important region in terms of both OGC standards development and the implementation of OGC standards in critical applications and portal applications.

In terms of standards development, for example, OGC Members from Taiwan submitted the candidate Open GeoSMS standard. This standards activity is also supported by organizations from Japan, Germany and the US. OGC Members from the Asia region are heavily involved in a variety of OGC Sensor Web Enablement (SWE) activities. Feng Chia University has signed a Memorandum of Understanding with the OGC to enhance and support our compliance testing capability for the OGC sensor web standards.

In terms of the implementation and use of OGC standards, the Asia region is extremely active. There are dozens of excellent examples. A few are described below.

- The International Centre for Integrated Mountain Development, a regional knowledge development and learning centre serving the eight regional member countries of the Hindu Kush-Himalayas - Afghanistan, Bangladesh, Bhutan, China, India, Myanmar, Nepal, and Pakistan has deployed a Mountain Geoportal (http://geoportal.icimod.org/) that implement the OGC Web Map Service (WMS) and KML standards as well as GeoRSS GML (Geography Markup Language).

- An activity known as Sensor Asia is Developing an infrastructure called Sensor Service Grid (SSG), which integrates field servers and Web GIS to realize easy and low cost installation and operation of ubiquitous field sensor networks. This application implements the OGC Sensor Observation Service (SOS), Observations and Measurements (O&M), and SensorML standards (http://www.mdpi.com/1424-8220/9/4/2363/pdf)
- The China Ministry of Land and Resources is using the OGC Web Feature Service (WFS) to build data exchange system. This system can implement the capability of two level (country level and province level) data exchange and satisfies the update requirements of 1:10000 land use status data.
- Ho Chi Minh City provides a tourism portal built using open source and OGC standards that focuses on the information needs of visitors and provides maps and information about things such as population, area, and education. A decision support system is being built on this framework, to serve tourism development and preservation of landscapes. JavaVietnam.org and CIREN Vietnam.

There are also numerous operational applications for emergency preparedness, alerting, and response. Several of these applications are described below.

Bryan Camoens: What are some of the challenges in formatting incoming data onto a common platform and how can you effectively overcome these challenges

Carl Reed: In this age of distributed computing, moving geospatial data from a geographically distributed set of databases into a single server or application tool no longer makes any sense for the majority of applications that require geospatial data (and services). The general industry consensus is that geospatial data, especially data being constantly updated, should remain as close to source as possible. The overhead of accessing and reformatting data for storage in a common system is a waste of time, effort and money. Further, the age of "static" formats such as SDTS (Spatial Data Transfer Standard) is over. Formats such as STDS are simply too inflexible, difficult to implement, and non-extensible.

Instead, systems are now accessing data from distributed sources in real time and applying semantic translations into a consistent data model or encoding and providing the data to the service or client that made the request. There are several technical reasons that this approach works effectively. First, communities of interest, or domains, are increasingly collaborating on a global basis to define content models with well-defined semantics. There are many examples, such as GeoSciML for sharing geological structure and bore hole data, WaterML for sharing water observation data, and CityGML for sharing 3d urban models. These content models have been developed by many organizations from many regions of the world, including Asia-Pacific. These content models can then be encoded using an XML-based language, such as the OGC/ISO Geography Markup Language (GML) or OGC/ISO Observations and Measurements.

The key aspect of the content modeling work is agreement on the semantics. This is the most difficult technical aspect of developing a consensus content model, especially on an international basis! By agreeing on the semantics (and vocabularies), mapping from an individual organization's database model to the consensus content model often becomes simply a matter of on-the-fly schema mapping. As an example, a client application can make a request to multiple, heterogeneous aeronautical chart data stores and receive content all in the same encoding (in this case AIXM/GML). There is no need for a common hardware/software platform. The common platform is now the consensus data model and the standards that support that model.

In terms of OGC developments, during the last two years there has been an increasing focus on domains, domain expertise, domain content models, and defining best practice use of OGC standards in those domains. The current active Domain Working Groups in the OGC are Aviation, 3d Information Management, Emergency and Disaster Management, Defense and Intelligence, Hydrology, as well as Meteorology and Oceans.

Lastly, the major impediments to data sharing are not technical – they are institutional, political and cultural. Having agreements between and among jurisdictions is critical to the ability share geospatial data, especially in time of need.

Bryan Camoens: When it comes to disaster and emergency management, how important is the use of geospatial tools and how should it be effectively utilized?

Carl Reed: OGC standards are widely used in emergency and disaster management applications in Asia. This includes applications for alerting and warning. Three such applications of note are the Taiwan Debris Flow Monitoring System, the German Indonesian Tsunami Early Warning System and the Japanese Disaster application.

Debris flows are a major issue in Taiwan. A debris flow is a fast moving mass of unconsolidated, saturated debris that looks like flowing concrete. They cause significant destruction and loss of life. The ability to quickly model the potential occurrence of a deadly debris flow and then warn citizens that may in the path of the flow is critical. There is now a nationwide operational monitoring, modeling and warning system. The advanced monitoring instruments include rain gauges, wire sensors, geophones, and charge-coupled device (CCD) cameras. The latest implementation of the system makes extensive use of the OGC Sensor Web Enablement standards. This enhancement has changed the way of collecting, fusing, and providing the debris flow data. Before implementation of the OGC sensor standards, observation data was burned to CD or sent by email to users.

The Boxing Day Tsunami 2004 triggered various international efforts focused on tsunami early warning for the Indian Ocean Basin. The activities resulted in a considerable progress in tsunami science, in particular concerning sensor systems and tsunami modelling. An early warning system architecture was specifically developed in a project called GITEWS (German Indonesian Tsunami Early Warning System). GITEWS was developed as a partnership between the German Government, the United Nations (UN), and a large number of Indonesian partners. GITEWS focuses on the upstream information from sensor systems to the warning center and on the delivery of reliable tsunami warning messages as fast as possible. The system requires the use of many different sensor systems. The different sensor systems deliver data in proprietary data formats and exhibit specific types of behavior. To utilize these sensors for decision support processes a flexible integration approach was developed based on the principles of a Service Oriented Architecture (SOA) and the specifications of the OGC Sensor Web Enablement Initiative (SWE) and the OASIS Common Alert Protocol standard.

GEOSS – the Global Earth Observation System of Systems – "seeks to connect the producers of environmental data and decision-support tools with the end users of these products, with the aim of enhancing the relevance of Earth observations to global issues." GEOSS has eight themes of societal benefit, including disaster management planning and response. The stated goal for the Disaster theme is, "The Global Earth Observation System of Systems is integrate Earth observations with other information to help planners reduce vulnerability, strengthen preparedness and early-warning measures and, after disaster strikes, rebuild housing and infrastructure in ways that limit future risks." Currently, thirteen countries from the Asia region participate in GEOSS. The definition and testing of the GEOSS architecture is being facilitated by the OGC (http://www.ogcnetwork.net/Aipilot). The implementation platform for GEOSS is standards based. Any organization that has data they wish to share via the GEOSS platform can do so by implementing the appropriate standards interfaces, such as the OGC Web Map or Web Feature Services. A list of registered components/services for disaster warning, management and planning can be found here: http://www.geoportal.org/web/guest/geo_search_overview?p_p_id=srgPortlet_WAR_geoportal&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=4&_srgPortlet_WAR_geoportal_searchType=browse&_srgPortlet_WAR_geoportal_sbaId=1

Bryan Camoens: What are some of the challenges in evaluating the collection of geospatial intelligence using Web 2.0, and what are some of the benefits of merging the two?

Carl Reed: By "Web 2.0" I am assuming you mean volunteered geographic information (VGI) such as OpenStreetMap, social network content that is geotagged in some way, sensor feeds, and so forth. This is a really good question. Last year, "big data" exceeded 1 zettabyte. That's 1 billion terabytes. Big data includes "web logs; RFID; sensor networks; social networks; Internet text and documents; Internet search indexing; call detail records; astronomy, atmospheric science, genomics, biogeochemical, biological, and other complex and/or interdisciplinary scientific research; military surveillance; medical records; photography archives; video archives; and large scale eCommerce" (Wikipedia, 2011). Big data introduces the requirement for fusion, fusion means, different types of fusion (data, object, sensor)…OGC is involved in fusion activities from OGC Web Services projects, OWS-7 in 2010 and OWS-8 now.

Many of these data sources represent potentially rich information for fusion with traditional GEOINT sources. A number of these sources, such as tweets and OpenStreetMap tend to be up-to-date, on the ground HUMINT. Others include up-to-date real-time sensor feeds, critical to GEOINT decision support. In all cases, these data sources provide an interesting set of issues when considering how and when to use them in some GEOINT fusion or decision support activity. The three main issues are:

- Provenance: this means the origin, or the source of something, or the history of the ownership or location of an object. When you get data from a national mapping agency or private sector content provider, you also get information on when the data were collected and compiled, the source of the data, who compiled the data, dates of the most recent updates, and why the data were collected. The same is not true of most Web 2.0 data. There is no consistent provenance information – if any.
- Data Quality: when individuals or organizations obtain digital map data from a mapping agency or private sector content provider, information on the quality of the data is provided. These formal mapping organizations can provide scale, measures of accuracy, what coordinate reference system the data are provided in and so on. They can also provide data quality information on the source data from which the map products were produced. The same is not true for much Web 2.0 social media or volunteered geographic data. Organizations and users need to be able to determine if a specific data source is "fit for purpose". These application dependent data use decisions require some level of information on data quality.
- Sheer Volume. A major concern when using big data is that processing such data sources is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time. This is obviously a major concern when lives are at stake! There is hope that cloud computing may provide the elastic and scalable computing platform necessary to process big data as required. Early indications are that with proper software architecture and implementation of processing algorithms that using the cloud can increase processing performance by an order of magnitude. However, the evidence also suggests that some processes are more amenable to performance enhancements using cloud computing.

The above issues are why organizations such as the OGC have active working groups addressing how to express data quality metadata in a standardized way or why the World Wide Web Consortium (W3C) is planning a new data provenance activity.

Bryan Camoens: Going forward how will geospatial intelligence evolve by the year 2020?

Carl Reed: Given the pace of change in the information technology realm, predicting how geospatial intelligence will evolve is not easy. That said, some observations:

- Sensors will become the primary source (95+ %) of all digital data used in geospatial intelligence applications. This means that there will be an increased requirement for hardware and software resources that can effectively access, task, and manage sensor networks. New systems will be required to process and effectively fuse sensor observations into GEOINT applications. Standard interfaces and encoding, such as the OGC Sensor Web Enablement suite of standards will be critical.

- Big Data will become a major source of both primary and ancillary location data for use in GEOINT applications. Hardware and software systems will evolve to efficiently capture, process, mine, and archive the massive amounts of data available. In a sense, much like the human brain, new algorithms will "scan" data streams and big data archives for content relevant to a given decision support application. Data of little or no interest to a given decision requirement will be ignored.
- New visualization technologies and processing algorithms will make full immersion, virtual (space and time) reality environments the norm for everyone in the GEOINT workflow, from back office to warfighters.
- New algorithms coupled with massive computing power will allow extremely fast and accurate pattern recognition – both human and digital. Whether the computing power will be in the cloud or not is problematic. Moore's law suggests that in the not too distant future a warfighter will be able to easily wear a supercomputer, perhaps one that is woven directly into their uniforms. Smart phones already have more computing power than the old Cray 1 supercomputer!
- Geospatial interface and encoding standards will be mandatory for any and all GEOINT application development and operational deployment.

These are just some observations. Woven into all of these possibilities are questions of security, fault tolerant access to data and computing resources, policy, privacy, and so forth.

Note:
HUMINT: Human intelligence
SIGINT: Signals intelligence
ELINT: Electronic intelligence
IMINT: Imagery intelligence
MASINT: Measurement and signature intelligence