International Journal of Spatial Data Infrastructures Research, 2012, Vol.7, 277-299

Interoperable Search Mechanisms for Web 2.0 Resources*

Laura Díaz[1], Manuela Núñez-Redó[1], David González[1], José Gil[1], Pau Aragó[1], Edward Pultar[1] and Joaquín Huerta[1]

[1]University Jaume I, GeoTech/INIT, Castelló de la Plana, Spain;
laura.diaz@uji.es; nunezm@uji.es; gonzaled@uji.es; jose.gil@uji.es; parago@uji.es; pultar@uji.es; huerta@uji.es

Abstract

We are currently witnessing ordinary citizens willing to share geospatial information using friendly and easy-to use tools provided by Web 2.0 platforms. These platforms act as social networks describing events with large social impacts. Social networks are filled with volunteered information before, during, and after events that occur near human settlements and urban areas. The amount of this geolocated information is increasing due to the increase of location-aware devices that allow users in the field to share knowledge about an event's evolution and impact. In order to retrieve this information one interacts with the different search mechanisms provided by various Web 2.0 services. This paper explores how to improve the interoperability of these various Web 2.0 platforms by providing a single service as a unique entry. This paper demonstrates the utility of the Open Geospatial Consortium's Open Search Geospatial and Time specification as an interface for a service that searches, retrieves and aggregates information available in different Web 2.0 services. We present how this information is useful in complementing other official and scientific information sources by providing an alternative, contemporary source of information. We demonstrate this with a proof of concept presented in a forest fire scenario. The intrinsic interoperability of the system is reflected in the collaborations shown with different information systems such as those at the biodiversity and forestry units in the Institute of Environment and Sustainability at the Joint Research Centre.

DOI: 10.2902/1725-0463.2012.07.art14

Keywords: Web 2.0, VGI, Citizen Science, Interoperability, Web Services, Open Search

1. INTRODUCTION

Geographical Information Systems (GIS) have become indispensable tools for organizing and exploiting the large amount of geospatial content for environmental sciences. Additionally, GIS provide a framework for multidisciplinary analysis (Ramamurthy, 2006).

Along with the evolution of other information systems, GIS has moved towards distributed environments based on Web Services and Service Oriented Architectures (SOA) (Aalst et al, 2007). In order to increase interoperability in the geospatial domain, the Open Geospatial Consortium (OGC) provides multiple standard specifications for data encodings and service interfaces (Percival, 2011). The combination of these standards allows for establishing Spatial Data Infrastructures (SDIs) (Masser, 2005). However, SDI top-down methodologies and complex publication mechanisms limit potential user contributions. Thus SDI suffer from a lackluster amount of user motivation regarding participation and content management (Coleman et al, 2009; Díaz et al, 2011). Recent natural disasters have demonstrated existing difficulties in accessing and efficiently exploiting geospatial resources in SDI (Zlatanova and Fabbri, 2009). The difficulties stem from the absence of sufficient available resources and a lack of collaboration between different geospatial infrastructures and components.

On the other hand, we are currently witnessing the consolidation of a new generation of the World Wide Web in which the key features are bottom-up methodologies and user participation. The Web is now a collaborative environment that has turned users into active providers (Alameh, 2003) providing a massive amount of information (Belimpasakis and Saaranen, 2010). This information is mostly georeferenced according to a user's location thus providing a large amount of geo-referenced information available in a wide variety of domains.

The integration of this Volunteered Geographic Information (VGI) provides a social view that complements scientific data. However, the integration of these data sources into SDIs poses new research challenges. For example, to retrieve VGI users must take into account the different capabilities and application programming interfaces (APIs) offered by each Web 2.0 service. To overcome this we propose an integrated, scalable solution that is based on standard specifications. The goal is to improve the interoperability between the many available heterogeneous Web 2.0 services. Hence this research addresses the following question:

How can we coalesce authoritative and volunteered citizen information to be utilized in environmental scenarios?

The goal is addressed with a middleware component that provides a single search interface. This aids in the discovery of information across different social networks and crowdsourced information platforms. The developed prototype is a discovery service that implements the OpenSearch Geo-Time standard interface specification (Gonçalves, 2011).

The remainder of this paper is structured as follows: Section 2 defines the background and related work including the Open Search specification. We present the software architecture of our approach in Section 3 and the prototype design in Section 4. Section 5 demonstrates the approach in a real world forest fire monitoring scenario and demonstrates the improvements gained by utilizing interoperable techniques in the integration of various other system components. The paper closes with discussions on the results and conclusions in Section 6.

2. BACKGROUND AND RELATED WORK

Geoscience research is an interdisciplinary field which benefits from the expertise of different specialists from disciplines such as remote sensing, biology, and technology (Goodchild, 2008). By means of SDI, these experts share and manage data in order to run scientific models and produce useful information about our environment. On the other hand, Web 2.0 Services and crowdsourcing platforms used by ordinary citizens are populated with spatial information (in varying spatio-temporal resolutions) offering a complementary vision to monitor our environment. In this section we discuss the sharing of geospatial content using standards-based and Web 2.0 approaches then examine how to leverage both information sources.

2.1. Geospatial Information Infrastructures

SDIs enable users to share geospatial content in a distributed manner following the SOA approach. They are widely known as facilitators to coordinate geospatial information (Dessers et al, 2011). Additionally, SDIs play a key role in supporting users and providers by giving them the ability to discover, visualize, and evaluate geospatial data at regional, national and global scales (Masser, 2005).

International initiatives such as the Global Earth Observation System of Systems (GEOSS) (Pearlman and Shibasaki, 2008) or the European Infrastructure (INSPIRE, 2007) describe the overall architecture and best practices for designing and implementing SDIs, where content is managed by means of regulated, standardized services. INSPIRE was adopted as a European directive in February 2007 and lays out a legal framework for the European SDI regarding policies and activities with environmental impact (INSPIRE, 2007). The technical level provides a range of interoperability standards available for integrating information systems (Mykkänen et al, 2008) and defining a network based on discovery, view, download, transformation, and invocation services. The inherent complexity of standardized SDI and the complex mechanisms of deployment get worse as SDIs grow (Béjar et al, 2009). In this way, the publication of content, associated traditionally with the providers, is an arduous task provoking a lack of updated content (Díaz et al, 2011).

2.2. Web 2.0 Services and Volunteered Geographic Information

The emergence of Web 2.0 platforms and easy to use tools encourages ordinary citizens to produce and share Geographic Information (GI) on the Internet. This is a different technique compared to the top-down building methodologies and complex publication mechanisms of SDIs. Web 2.0-based activities show that users are willing to engage more actively in content creation. The progress in information technology and the conversion of the user into a provider have created the phenomenon and resulting concepts such as Web 2.0, Neogeography (Turner, 2006), Cybercartography (Tulloch, 2007) and volunteered geographic information (VGI) (Goodchild, 2007).

VGI, as an alternative source of information, to complement official information (Craglia et al, 2008), could be integrated within the SDIs context in order to improve traditional geospatial analysis and decision support tasks (Flanagin and Metzger, 2008; Pultar et al, 2009). Zook et al (2010) pointed out that VGI could provide "additional data at levels of granularity and timeliness that could not be matched by other means" (p. nr. 12).

Hybrid approaches for the integration of top-down and bottom-up methodologies to study the integration of Web 2.0 resources within the SDI context are founded on what has been coined as the 'reconceptualization of the SDI user role' (Budhathoki et al, 2008; Omran and van Etten, 2007; Goodchild, 2010). Here the new SDI generation will be influenced by more active user participation. Other researchers have paid attention to the versatility of Web 2.0 systems in contrast to SDI maintenance and publication mechanisms, aiming to lower the barrier in the SDI publication mechanism (Díaz et al, 2011). Recent research has also studied the retrieval of data directly from crowdsourcing services where a discovery service deployed on an INSPIRE-based infrastructure offers a standard, unique entry point for VGI retrieval (Nuñez et al, 2011a).

2.3. Geospatial Content Discovery

Within SDIs, metadata and catalog services are key for properly discovering content (Craglia et al, 2007). The services support the ability to publish and search for metadata while also supporting information resource binding within SDI. In this context, Bigagli et al (2004) proposed GI-Cat, a SOAP-based Web service providing basic functionalities of geospatial information cataloguing.

The EuroGEOSS Discovery Broker (Nativi et al, 2011) component based on GI-Cat is used in the context of the European Union project EuroGEOSS[1]. This allows for a unique access point to services provided by three disciplines covered in the project: biodiversity, forestry, and drought.

Web 2.0 services are immense online repositories with geo-referenced content. However, attempts at providing spatio-temporal search engines for VGI are relatively scarce (Tsai, 2011). The process of searching through multiple services becomes a tedious task because each service provides different data encodings, geo-referencing, and proprietary APIs. It is a goal of this work to provide an interoperable mechanism to search Web 2.0 content to be integrated into a SDI context. Within the EuroGEOSS project the results of this research will be used to provide the EuroGEOSS Discovery Broker with Web 2.0 services as an additional source of content for the SDI users.

2.4. The OGC Open Search Geo-Time Specification

Web 2.0 services provide public APIs which allow custom programs to interact with the services via specific encodings and functions. Since there is no standard set of methods or encodings, there is a technical barrier for discovering content from multiple sources in a homogeneous way.

The Open Geospatial Consortium (OGC) OpenSearch Geospatial and Time specification (OSGT) (OGC, 2010) describes an interface based on minimal mandatory input which can be extended with spatial and temporal criteria. OpenSearch's simplicity has made it rapidly become a successful search interface specification for Web repositories being used by sites such as Flickr and Wikipedia. This research has adopted the interface of the OSGT specification as further described in the following sections.

OpenSearch defines a service interface for minimal search and retrieval capabilities. An OpenSearch-enabled service exposes an interface for client applications to send simple HTTP GET requests providing specific query parameters. As a result, responses are often encoded in lightweight data formats

[1] http://www.eurogeoss.eu

such as GeoRSS[2], Atom[3] or KML[4]. Each service must be described by its Description Document. This document is a file that describes the search engine of the target service. The description may vary from one service to another, but there are several mandatory parameters:

- Root node called OpenSearchDescription;
- ShortName which contains a brief human-readable title to identify the search engine;
- Description which is a text description of the search engine;
- URL with the location of where a search request may be executed.

The OpenSearch specification has only one mandatory query parameter called “searchTerms” allowing client applications to request information related to one or more keywords. Other query parameters such as those supporting pagination of results (i.e. “count”, “startIndex”, “startPage”) are optional.

Specific search profiles are described by extending the OpenSearch specification. The OGC OpenSearch Geo Temporal specification defines a list of parameters to enable spatial and temporal filtering. This specification profile allows the user to filter results by a particular place name, area, point and radius and by time period.

3. SYSTEM ARCHITECTURE

In this section we elaborate on the architecture of the proposed approach. Our main goal is to extend traditional SDI architecture with a middleware component that offers a standard interface to retrieve and integrate Web 2.0 content and information from SDI sources.

Figure 1 shows a simplified overview of the proposed architecture based on INSPIRE. This architecture is composed (top-down) of the application layer, service layer, and content layer. SDIs based on INSPIRE provide functionality to users by means of categorized services deployed on the service layer. Following these principles and the goal of having an interoperable way to access and integrate Web 2.0 content in the SDI context, we propose to extend this architecture with a new Discovery Service. This new Discovery Service will provide the capability of Web 2.0 content retrieval and integration.

---

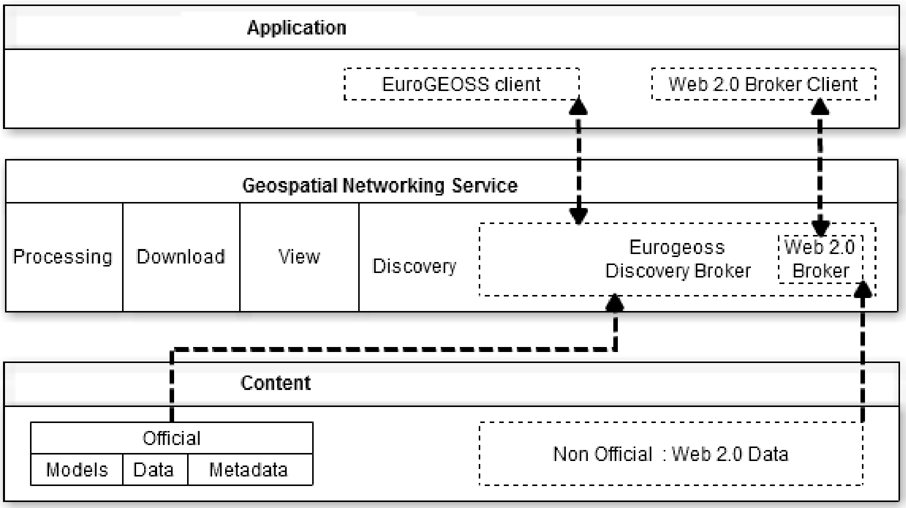
[2] http://www.georss.org
[3] http://tools.ietf.org/html/rfc4287
[4] http://www.oprngeospatial.org/standards/kml

3.1. Content Layer

We focus on the integration of both official and unofficial sources. Specifically, this work describes the retrieval of content provided by Web 2.0 services. These unofficial resources are especially relevant due to the fact that users provide near real-time information and local knowledge that enriches official information.

Due to the substantial availability of resources in crowdsourcing platforms a large part of the retrieved results are not related to the target scenario and they present "noise" that has to be eliminated for an appropriate assessment. This is a relevant issue that requires further investigation of the retrieved content and its semantics. Although this is out of the scope of this work it is considered one of the most promising future lines of work.

Figure 1: Three-Tier SDI Architecture with a New Connection to VGI Sources



3.2. Geospatial Networking Service Layer

A classic SDI provides discovery, view, download and processing services that implement the standards-based interfaces in order to improve the interoperability of the system and their components. In our research we propose to extend this layer by adding a new discovery service called the Web 2.0 Broker (W2.0B). Following a brokering approach, the W2.0B provides the capacity to perform a spatio-temporal search within multiple Web 2.0 services and offers a homogeneus, unique entry point. To do this we have our W2.0B discovery service implement the OGC OSGT interface. For a better integration of crowdsourced data with official sources, such that they can be crossed and represented together, we have experimented with the integration of the W2.0B component as part of the EuroGEOSS Discovery Broker to create a search engine for accessing Web 2.0 content.

The results of the Web 2.0 search can then be integrated with official information available in the context of an SDI client. The client application is able to integrate and visualize this information retrieved from the W2.0B which provides an aggregated list with the resources in a common format. As a result the vast amount of VGI becomes a new source of information to complement scientific data.

3.3. Application Layer

In this layer users are presented with user interfaces, such as geoportals, providing the entry point to the functionality offered by the services. In this particular case, to demonstrate the functionality and added value of the Web2.0B, we deploy and test two types of Geoportals that access this service. The first is a dedicated Web 2.0 client application with an intuitive interface to perform queries according to keywords and spatio-temporal criteria. The W2.0B users may query to retrieve results that will be transformed and presented in this layer using Web mapping technology. Second, existing SDIs and Geoportals (within the EuroGEOSS project) will make use of the functionality provided by the W2.0B both directly or indirectly through the EuroGEOSS Discovery Broker.

4. IMPROVING WEB 2.0 SERVICES INTEROPERABILITY: THE WEB 2.0 BROKER

One main goal is to improve interoperability when accessing the multitude of different APIs published by Web 2.0 Services. The W2.0B follows a brokering approach and implements the Open Search Geo-Time specification to provide the ability to search and retrieve content from different Web 2.0 Services. A collection of social media services with geo-referencing capabilities have been analyzed in Table 1. Only those that support geospatial and temporal filtering functions through their public API have been selected as target repositories (Fonts et al, 2010). Although some Web 2.0 services implement the OpenSearch specification (e.g. Flickr, Wikipedia, Youtube), some of them do not offer the OpenSearch Geo-Time search interface. The W2.0B overcomes this limitation by offering spatial and temporal criteria queries to the services that provide this information.

Table 1 shows the current status of the W2.0B prototype. The colored cells show the services currently offered where the rows show both the GI resources and the Web 2.0 services. The columns show the operations available in the OSGT specification and implemented in order to query the Web 2.0 services. Results are available in Atom, KML and MIMETEXT KML (Abargues et al, 2010) formats.

4.1. Web 2.0 Broker –Design

Figure 2 shows the component diagram of the W2.0B. It illustrates its modular design, its workflow and the placement of its components. At the top of Figure 2 one can see how the W2.0B implements the OSGT specification.

The OpenSearch *Descriptor Manager* (OS Descriptor Manager) dynamically generates the service description document[5]. This is a mandatory document specifying and detailing the service capabilities. It provides information about the valid query parameters and supported response formats. This allows the client application to understand the interfaces needed to build valid OpenSearch-style queries.

This component also allows for integrating a custom search engine within web browsers such as Internet Explorer, Firefox, and Opera, allowing users to run text-based searches in the Web2.0B.

The *OpenSearch core (OS Core)* component deals with the interpretation of the query in the standard OpenSearch format. It retrieves the query and forwards it to the *Search Engine Manager* component. Only the SearchTerms input parameter from the specification is mandatory, but other criteria can be specified. The accuracy of results is improved by adding spatial or temporal criteria.

The Search Engine Manager component is in charge of transforming the query and the specified criteria to the concrete Web 2.0 services APIs. It plays a mediating role between service-specific APIs and the OpenSearch query by broadcasting the query to the requested Web 2.0 services. The capabilities supported by W2.0B are limited to the native functionality offered by each specific API.

However, not all of the Web 2.0 services have an API for accessing data. For instance, weather information extracted from Meteoclimatic or environmental news from European media sites (EMM) are provided as a data stream that is interpreted and transformed into a proper data format. This means that users cannot perform custom queries. Currently the information retrieved from this source is parsed to extract data related to fire news in order to assess the value of this data in our forest fire monitoring use case.

The *Geo-reference Manager* component manages and improves the geocoding of the retrieved content in terms of spatial search accuracy and performance. This is critical since some results provided by the services do not contain any location. The core of this module,concerns the improvement of un-georeferenced

[5] http://www.openseach.org

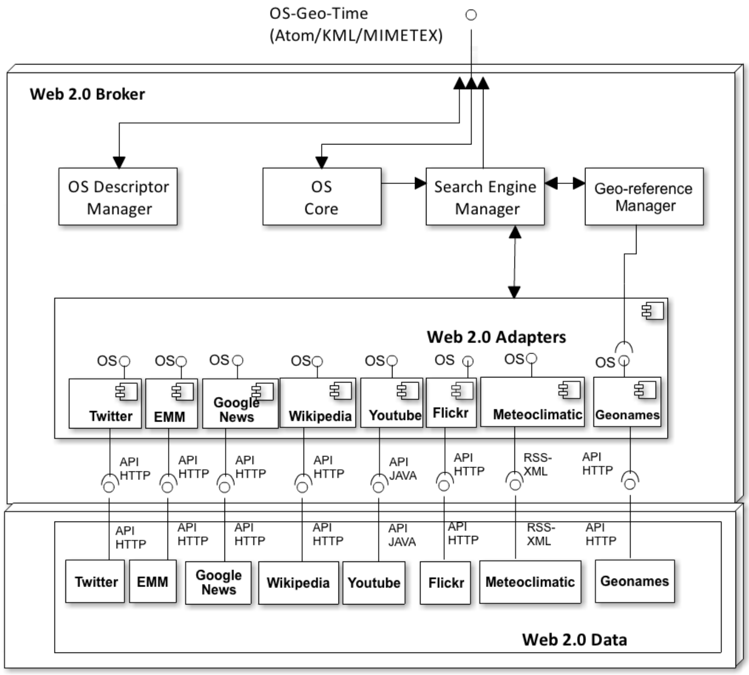
data by connecting to geonames[6] to extract coordinates from placenames enhancing the location of the result.

Table 1: Search Parameters Implemented in the W2.0B Applicable to Web 2.0 Services

		Base OS params				Geo OS Extension				Time Ext		Data Formats		
		Search Terms	Count	Start Index	Start Page	Bbox	Lon,Lat,Rad	Geometry	Name/Location	Start	End	KML	KML/MIMEXT	ATOM(GeoRSS)
Web 2.0 Services	Twitter	■	■		■		■			■	■	■	■	■
	Open Street Map	■				■			■			■	■	■
	Meteoclimatic	■				■						■	■	■
	Youtube	■					■			■	■	■	■	■
	Flickr	■	■		■	■				■	■	■	■	■
	Geonames	■							■			■	■	■
	Wikipedia	■										■	■	■
	Google News	■								■	■	■	■	■
	EMM	■								■	■	■	■	■
Catalogue Services	Eurogeoss Discovery Broker	■	■	■		■	■			■	■	■	■	■

[6] http://www.geonames.org/

Figure 2: W2.0B Components Diagram.



This broker architecture is flexible enough to add new Web 2.0 services to the system without altering the broker's discovery interface from the perspective of the client. Thus clients and modules are independent, loosely coupled components where each one evolves separately thereby enhancing the system scalability as a whole (Nuñez et al, 2011b).

4.2. Web 2.0 Broker Implementation

The W2.0B has been designed as a service with a standard interface to be re-used in different scenarios. In this section we illustrate how the W2.0B works when it is invoked.

Figure 3 illustrates the workflow of how the different components work together when the user intitiates a query. As an example, we propose a query to search for photos related to the term "incendio" in the area of spain.

During the first step the client application wil retrieve the Flickr descriptor via the OpenSearch Descriptor Manager Component to see the service capacities in terms of search criteria and data encodings.

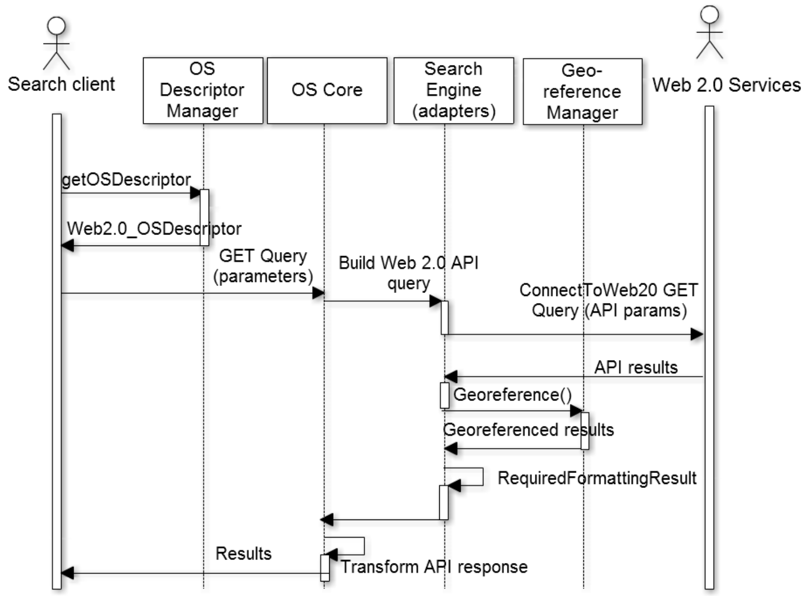
At this point, the client is able to build and send a well-formatted OpenSearch query:

```
http://
elcano.dlsi.uji.es/WB20_BROKER2/broker.jsp?service=fck&q=incendio&bbox=-
4.456,36.752,-4.328,36.809&lon=-
3.757&lat=40.048&radius=1040.122&loc=&start=&end=&format=kml).
```

When the W2.0B receives the query, the OpenSearch core interprets it. The Search Engine contains a collection of Search engines or adapters that transform the query to a specific Web 2.0 API style. Next the W2.0B propagates the query to the different services by means of each search engine.

Since many of the retrieved results might not contain any location information, the next step is to improve georeferencing. The Geo-Reference component is in charge of parsing the result to find information as a toponym and using third party services such as Geonames. The retrieved results are aggregated and returned in a response encoded in the format requested by the client.

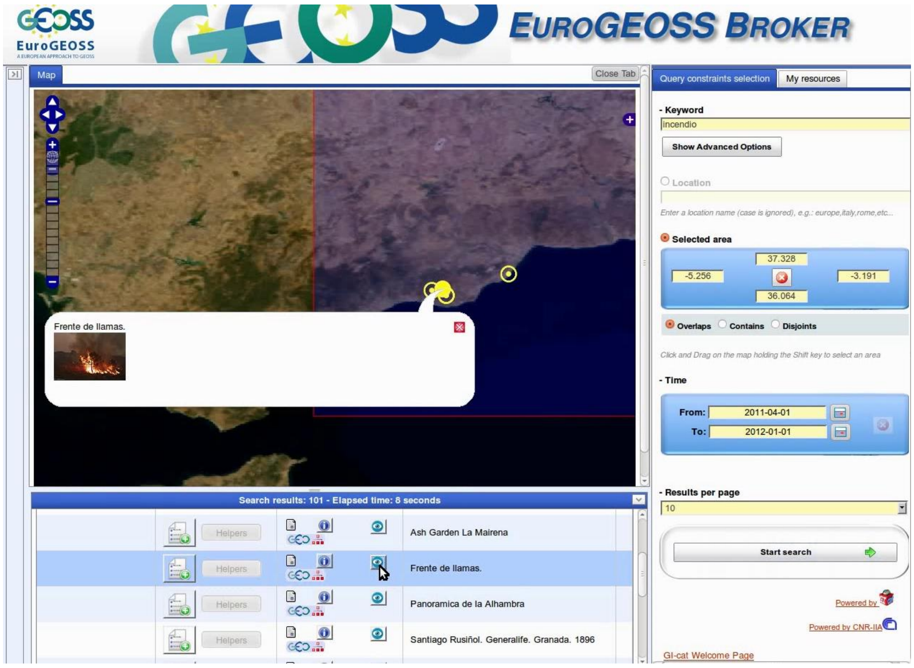
Figure 3: W2.0B Sequence Diagram for Performing a Query



4.3. Integration of Web 2.0 Broker into EuroGEOSS Brokering Platform

The EuroGEOSS broker provides multiple profilers (different interfaces to query data) and accessors (able to integrate multiple and heterogeneus sources) to increase interoperability. W2.0B is integrated into the EuroGEOSS brokering platform as an accessor of the Discovery Broker. Its fundamental aim is retrieving information from social networks which will be aggregated and returned to the client. The W2.0B acting as a interoperable accessor and enables Web 2.0 services to be accessed through all of the protocols available in the EuroGEOSS broker's profilers components. Figure 4 presents a screenshot of the EuroGEOSS Broker client[7] demonstrating how this integration retrieves Web 2.0 data and shows VGI in the form of a photograph retrieved from Flickr.

Figure 4: Screenshot of W2.0B Integration into Eurogeoss Brokering Platform



5. MULTIDISCIPLINARY ENVIRONMENTAL MONITORING SCENARIO: FOREST FIRE MONITORING AND IMPACT ASSESSMENT

Environmental monitoring is a discipline that involves many possible scenarios with multiple variables and procedures. To illustrate the practical use of the W2.0B prototype we describe a multidisciplinary scenario involving forest fire monitoring. During and after the fire, environmental experts must monitor the area in order to evaluate the fire damage as well as environmental (forestry resource, biodiversity loss and drought influences) and social impacts.

The monitoring during the post-fire phase will use both citizen science by collecting volunteered contributions through Web 2.0 services in addition to

[7] http://www.eurogeoss-broker.eu/

official environmental-related data. We use a real world scenario to demonstrate the utility of the W2.0B: a fire event that occurred in Mijas, Málaga, Spain in 2011.

In order to monitor the status of a detected fire, the user accesses the Web2.0B client (see Figure 1) which provides a web map to visualize data coming from various sources. Additionally, the client shows the burned area retrieved from the Europen Forest Fire Information System (EFFIS) data services[8] which allows one to see official sources in addition to the unofficial VGI.

This web client offers simple and advanced interfaces (customizable by the user) to specify search criteria and build a query. Users can add spatio-temporal criteria by selecting the area of interest in the form of a rectangle on the map or by providing point and radius information. Additionally, users looking for results within a certain time period may specify begin and end dates. The client is currently available to the public at http://geoinfo.dlsi.uji.es/GF/.

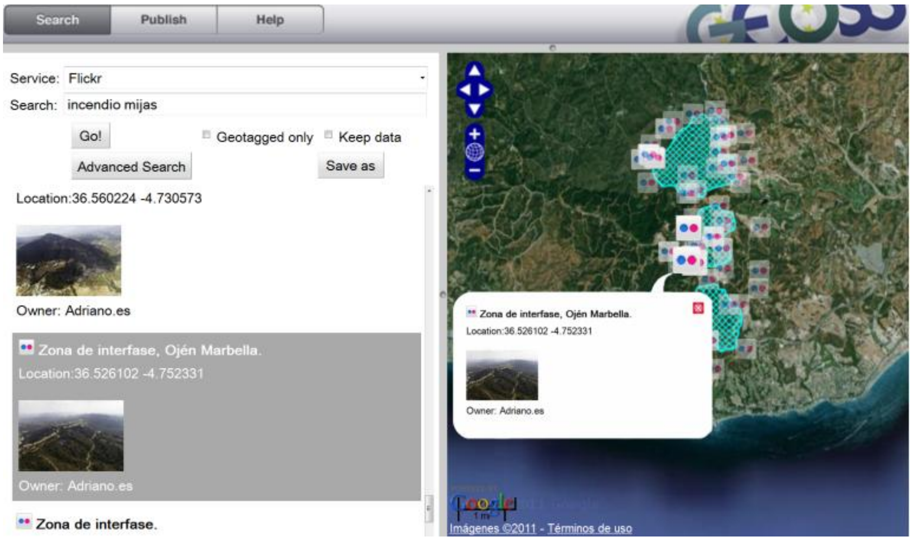
For this example in Mijas, Spain, the user restricts the area of interest to Mijas and limits the time period in order to retrieve Web 2.0 information of interest. The client application generates the OSGT query and connects to the W2.0B to retrieve information with the spatio-temporal criteria as shown in Figure 5. The information retrieved here explicitly demonstrates the power of integrating authoritative, official data with citizen-based, unofficial Web 2.0 data. In particular, Figure 5 shows pictures taken from users who were near the forest fire when it happened. In the figure one sees the added value of the citizen-based, Web 2.0 data and how it complements official data sources. The convex hull of the pictures provided by the citizens closely approximates the official burned area provided by EFFIS. Here in the absense of official burn area the citizen data gives a worthy estimate. In addition, the volunteered Web 2.0 data is available in near real time meaning that in most cases this information will be useful for time critical decisions thereby saving more human lives by acting on this data immediately rather than waiting for official sources.

One of the main objectives of the W2.0B is to aid in studies by offering a citizen’s point of view since citizens contribute an alternative source of information. The VGI in Figure 5 extracted from Flickr presents the Wildland-Urban Interface (WUI) (Stewart et al, 2007) between Ojén and Marbella in Spain. It has been demonstrated that identifying the area where houses meet or intermingle with undeveloped wild land vegetation will contribute to preventing and fighting forest fires (Haight, 2004). Additionally, if this information is collected in real time it will aid in dynamic GIS analysis for when to evacuate a community in the event of a wildfire (Pultar et al, 2009).

[8] effis.jrc.ec.europa.eu

VGI is inherently multidisciplinar, users provide information about a wide range of discilines. In ur study case when retrieven data from the fire and its evolution we can also search and retrieve information regarding biodiversity to know which type of flora and which species are being affected by the fire in Mijas.

Figure 5: Screenshot of W2.0b Results Retrieved from Flickr when Doing a Search for "Incendio Mijas"and Burned Areas from EFFIS through the Eurogeoss Discovery Broker



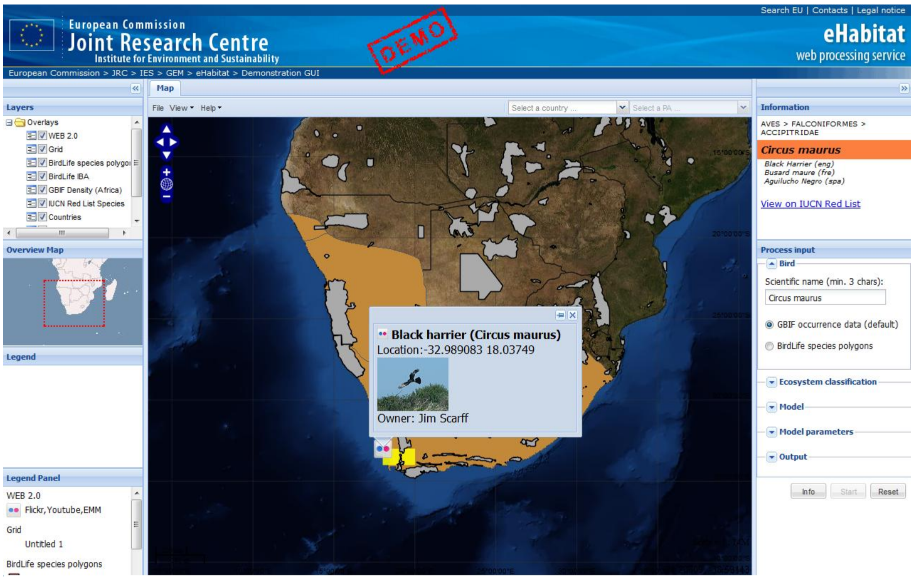
W2.0B is a tool for retrieveing data to assist in multidisciplinary environmental monitoring based on citizen contributions. This is not a substitute for scientific data but is a complementary source of information for natural disasters and hazards such as wildfires or hurricanes. Further steps are needed to analyze and model raw VGI in order to extract more accurate and relevant information out of the massive repository provided by social networks.

The W2.0B component is based on standard specifications which improves its ability to work with other components and systems. This aspect of the component is illustrated in the following examples where W2.0B's added value is shown through its integration into third party applications such as eHabitat (Dubois et al, 2011) and geo-wiki (Fritz et al, 2009). Figure 6 shows the use of the W2.0B component to retrieve information about a bird called "Circus Maurus" in South Africa through the eHabitat[9] platform. This platform is designed for finding and assessing ecosystems with equal properties by the biodiversity department at the

[9] http://ehabitat-wps.jrc.ec.europa.eu/ehabitat/

Institute of Environment and Sustainability (IES) at the Joint Research Center (JRC).

Figure 6: Screenshot of eHabitat.



Also, Figure 7 demonstrates how the forestry department at the IES of the JRC integrated a W2.0B client[10] to access the features of this component for integrating scientific and citizen-based data into their system (McInerney et al, 2012).

Figure 7: Screenshot of Third Partie Eurogeoss Forestry Web

[10] http://193.126.113.48/wp3jrc/faoc/ext_view.html

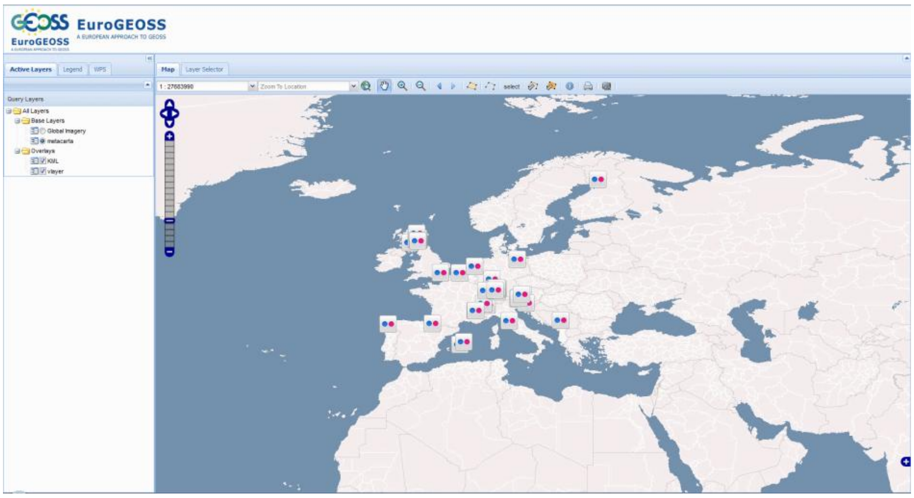
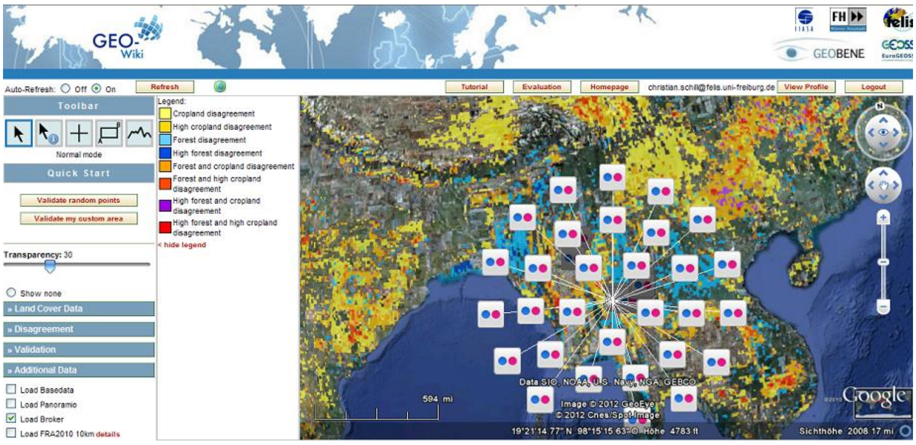


Figure 8 shows results obtained from the W2.0B integrated in the geo-wiki[11] project, which is a global network of volunteers aiming to improve the quality of different thematic datasets. In this example the VGI extracted from Web 2.0 services is related to Flickr content tagged as "nature" in Thailand (Fritz et al, 2012).

Figure 8: Screenshot of Integrating the W2.0B in the Geo-Wiki Platform



6. DISCUSSION AND CONCLUSIONS

The increasing amount of information provided by citizens in crowdsourcing platforms should not be ignored. This information is spatially referenced in a

[11] http://geo-wiki.org

small percentage but this percentage is also growing together with the use of sensor-enabled devices such a smartphones. During this work we address the need to assess what could we do with the information already available and referenced.

Due mostly to heterogeneity, leveraging this information presents many challenges in order to increase interoperability and scalability as well as integrate and consume them in a scientific context. In this line future research to assess aspects like quality (Ostermann and Spinsanti, 2011) have to be considered and integrated.

Following the SDI principles we have proposed a new Discovery Service which provides a standards-based, unique entry point to query multiple Web 2.0 services and retrieve citizen-based information to be integrated in SDI contexts. This mechanism, the Web 2.0 Broker, interprets queries based on the standards-based Open Search Geo-Time specification.

The Web 2.0 Broker is able to interpret OSGT queries and propagate them to a set of Web 2.0 Services then aggregate the results encoded in standard data formats such as GeoRSS, GeoJSON, KML, or ATOM.

One advantage of the proposed approach is that potential calibration mechanisms can be encapsulated in well-defined components that directly connect and use the specific Web 2.0 service APIs. However, the search criteria based on OpenSearch need to be mapped into the specific Web 2.0 service APIs and this means that we could lose accuracy in certain parts of a query. This may have an impact on the number of VGI items retrieved and therefore further work and analysis is needed in this respect.

Our work indicates that VGI can complement SDI data by providing high-scale, value-added information at a low cost. In our study case we have demostrate how the retrieved VGI allow users to monitor the evolution of the fire. Figure 5 is a representative example, where Flickr pictures (with spatial information) are place on top of the official data: satellite image and burned area retrieved by the EFFIS. We can see how the pictures are a good aproximation of the burned area, which could have been aproximately calculated using this VGI. In this direction, the W2.0B can be used to complement autoritathive data in may scenarios such as crisis management. The use case scenario presented in this paper demonstrated the value added by our multidisciplinary approach. The scenario illustrates how general environmental monitoring and particularly forest fire monitoring can leverage the potential of available VGI (Nunez et al, 2011). In our case we have consumed raw VGI after performing a preliminary visual analysis. The validation of global models with this knowledge extracted by VGI analysis has yet to be fully exploited. There is a need for more sophisticated analysis to filter the massive amount of data, to extract more accurate information and avoid the inherent noise of consuming raw data (especially Web 2.0 data). Modelling this data in order to detect specific patterns and changes can generate more relevant and accurate information.

Further development of the Web 2.0 Broker is currently on-going. We aim to improve the means by which VGI is harnessed and integrated into SDIs thus leveraging its full potential. This will be done by increasing the number of Web 2.0 resources to be aggregated and by analyzing the data flow to better extract observations relevant to specific use cases. The next steps are to define a data model to describe forest fire observations and alarms in order to add a new source of information for emergency response scenarios. Additionally, we are exploring VGI data mining to infer and assess green routes for environmental management as well as extract meaningful places for urban management .The intrinsic multidisciplinary character of the W2.0B component favors the utility in a wide variety of use cases that will carry on well into the future.

ACKNOWLEDGEMENTS

This work has been partially supported by the European FP7 Project nr. 226487 called EuroGEOSS.

REFERENCES

Aalst, W.M.P. van der, Beisiegel, M., Hee, K.M. van der, Konig, D. and C. Stahl (2007). An SOA-based architecture framework, *International Journal of Business Process Integration and Management,* 2(2): 91-101.

Abargues, C., Beltrán, A. and C. Granell (2010). “MIMEXT: a KML Extension for Georeferencing and Easy Share MIME Type Resources”, in Painho, M., Santos, M.Y. and H. Pundt (Eds.) *Geospatial Thinking*, Berlin, Heidelberg: Springer, pp. 315-334.

Alameh, N. (2003). Service chaining of interoperable geographic information web services, *IEEE Internet Computing*, 7(5): 22-29.

Béjar, R., Latre, M.Á., Nogueras-Iso, J., Muro-Medrano, P.R. and F.J. Zarazaga-Soria (2009). An architectural style for Spatial Data Infrastructures, *International Journal of Geographical Information Science*, 23(3): 271-294.

Belimpasakis, P. and A. Saaranen (2010). Sharing with people: a system for user-centric content sharing, *Multimedia Systems,* 16(6): 399-421.

Bigagli, L., Nativi, S., Mazzetti, P. and G. Villoresi (2004). ”GI-Cat: a Web service for dataset cataloguing based on ISO 19115”, *Proceedings of the 15th*

International Database and Expert Systems Applications Workshop (DEXA'04), August 30-September 3 2004, Zaragoza, Spain, pp. 846-850.

Budhathoki, N.R., Bertram, B. and Z. Nedovic-Budic (2008). Reconceptualizing the role of the user of spatial data infrastructure, *GeoJournal,* 72(3): 149-160.

Coleman, D.J., Georgiadou, P.Y. and J. Labonte (2009). Volunteered geographic information: the nature and motivation of produsers, *International Journal of Spatial Data Infrastructures Research*, 4: 332-358.

Craglia, M. (2007). “Volunteered Geographic Information and Spatial Data Infrastructures: when do parallel lines converge?”, *Position paper for the VGI Specialist Meeting, December 13-14 2007, Santa Barbara, USA,* at http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Craglia_paper.pdf [accessed 14 September 2012].

Craglia, M., Kanellopoulos, I. and P. Smits (2007). ”Metadata: where we are now, and where we should be going”, *Proceedings of the 10th AGILE International Conference on Geographic Information Science 2007, May 8-11 2007, Aalborg University, Denmark.*

Craglia, M., Goodchild, M.F., Annoni, A., Câmara, G., Gould, M., Kuhn, W., Mark, D., Masser, I., Maguire, D., Liang, S. and E. Parsons (2008). Next-Generation Digital Earth: A position paper from the Vespucci Initiative for the Advancement of Geographic Information Science, *International Journal of Spatial Data Infrastructures Research,* 3: 146-167.

Dessers, E., Crompvoets, J., Janssen, K., Vancauwenberghe, G., Vandenbroucke, D., Vanhaverbeke, L. and G.A. Van Hootegem (2011). “Analysing spatially enabled work processes. Results of a multidisciplinary case study on spatial planning”, *Pre-Conference Workshop “Multi - and Interdisciplinary Research on Spatial Knowledge in the light of SII”, AGILE International Conference on Geographic Information Science, April 18-22, 2011, Utrecht, The Netherlands*.

Díaz, L., Granell, C., Gould, M. and J. Huerta (2011). Managing user generated information in geospatial cyberinfrastructures, *Future Generation Computer Systems*, 27(3): 304-314.

Dubois, G., Skoien, J., Mendes de jesus, J., Peedell, S., Hartley, A., Nativi, S., Santoro, M. and G. Geller (2011). “eHabitat: A Contribution to the Model Web for Habitat Assessments and Ecological Forecasting”, *Proceedings of the 34th International Symposium on Remote Sensing of Environment – The GEOSS Era: Towards Operational Environmental Monitoring, April 10-15 2011, Sydney, Australia,* at http://www.isprs.org/proceedings/2011/ISRSE-34/211104015Final00047.pdf [accessed 14 September 2012].

Flanagin, A.J. and M.J. Metzger (2008). The credibility of volunteered geographic information, *GeoJournal*, 72(3): 137-148.

Fonts, O., Huerta, J., Díaz, L. and C. Granell (2010). “OpenSearch-geo: The simple standard for geographic search engines”, *Proceedings of IV Jornadas SIG Libre, March 10-12 2010, Girona, Spain,* at http://dugi-doc.udg.edu/bitstream/handle/10256/3507/a29.pdf?sequence=1 [accessed 14 September 2012].

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F. and M. Obersteiner (2009). Geo-Wiki.Org: The Use of Crowdsourcing, *Remote Sensing*, 1(3): 345-354.

Fritz S., McCallum I., Schill C., Perger C., See L., Schepaschenko D., van der Velde M., Kraxner F. and M. Obersteiner (2012). Geo-Wiki: An online platform for improving global land cover, *Environmental Modelling & Software,* 31: 110-123.

Gonçalves, P. (Editor) (2011) OpenGIS® OpenSearch Geospatial Extensions Draft Implementation Standard, Version 0.0.3, Open Geospatial Consortium Inc. Ref. OGC 10-032r3, at http://wiki.services.eoportal.org/tiki-download_wiki_attachment.php?attId=1447 [accessed 14 September 2012].

Goodchild, M.F. (2007). Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0, *International Journal of Spatial Data Infrastructures Research*, 2: 24-32.

Goodchild, M.F. (2008). NeoGeography and the nature of geographic expertise, *Journal of Location Based Services*, 3(2): 82-96.

Goodchild, M.F. (2010). Twenty years of progress: GIScience in 2010, *Journal of Spatial Information Science*, 1: 3-20.

Haight, R.G., Cleland, D.T., Hammer, R.B., Radeloff, V.C. and T.S. Rupp (2004). Assessing Fire Risk in Wildland-Urban Interface, *Journal of Forestry*, 102(7): 41-48.

INSPIRE EU Directive, (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Official Journal of the European Union, L 108/1, Volume 50.

Masser, I. (2005). *GIS Worlds: Creating Spatial Data Infrastructures, Redlands: ESRI Press*.

McInerney, D., Revez, G., Díaz, L., Figueirido, C., Barredo, J.I. and J. San-Miguel-Ayanz (2012). "Web-based Forest Monitoring tools using OGC Services*", EuroGEOSS Conference, January 25-27 2012, Madrid, Spain.*

Mykkänen, J. and M. Tuomainen (2008). An evaluation and selection framework for interoperability standards*, Information and Software Technology*, 50(3): 176-197.

Nativi, S., Craglia, M., Vaccari, L. and M. Santoro (2011). "Searching the New Grail: Inter-Disciplinary Interoperability"*, Proceedings of the 14th AGILE International Conference on Geographic Information Science 2011, April 18-22 2011, Utrecht, The Netherlands.*

Nuñez, M., Díaz, L., Granell, C. and J. Huerta (2011a). "Web 2.0 Broker: a tool for massive collection of user information", *Geophysical Research Abstracts, European Geosciences Union General Assembly, April 3-8 2011, Vienna, Austria.*

Nuñez, M., Díaz, L., Gil, J., González, D. and J. Huerta (2011b). "Discovery and Integration of Web 2.0 Content into Geospatial Information Infrastuctures: A use case un Wild Fire Monitoring", *Proceedings of the IFIP WG 8.4/8.9 international cross domain conference on Availability, reliability and security for business, enterprise and health information systems (ARES'11).* Berlin, Heidelberg: Springer-Verlag, pp. 50-68.

Omran, E.L.E. and J. van Etten (2007). Spatial-data sharing: Applying social-network analysis to study individual and collective behaviour. *International Journal of Geographical Information Science*, 21(6): 699-714.

Ostermann, F. and L. Spinsanti (2011). "A Conceptual Workflow For Automatically Assessing The Quality Of Volunteered Geographic Information For Crisis Management," *Proceedings of the 14th AGILE International Conference on Geographic Information Science, April 18-22 2011, Utrecht, The Netherlands*.

Pearlman, J. and R. Shibasaki (2008). Guest Editorial, *IEEE Systems Journal*, 2(3): 302-303.

Percival, G. (Editor) (2011) OGC Reference Model (ORM) version 2.1. Open Geospatial Consortium. OGC 08-062r7, at https://portal.opengeospatial.org/files/?artifact_id=47245 [accessed 14 September 2012].

Pultar, E., Raubal, M., Cova, T.J. and M.F. Goodchild (2009). Dynamic GIS Case Studies: Wildfire Evacuation and Volunteered Geographic Information, *Transactions in GIS*, 13: 85-104.

Ramamurthy, M.K (2006). A new generation of cyberinfrastructure and data services for earth science education and research, *Advances in Geosciences,* 8: 69-78.

Stewart S.I., Radeloff, V.C., Hammer, R.B. and T.J. Hawbaker (2007). Defining the Wildland–urban Interface, *Journal of Forestry*, 105(4): 201-207.

Tsai, F.S. (2011). Web-based geographic search engine for location-aware search in Singapore, *Expert Systems with Applications*, 38(1): 1011-1016.

Tulloch, D.L. (2007). Many many maps: Empowerment and online participatory mapping. *First Monday*, 12(2).

Turner, A. (2006). *Introduction to Neogeography*, O'Reilly Media Inc.

Zlatanova, S. and A.G. Fabbri (2009). Geo-ICT for Risk and Disaster Management, in Scholten H.J, Van de Velde R. and & van Manen N. (Eds) *Geospatial Technology and the Role of Locations in Science*, Dordrecht, Springer, pp. 239-266.

Zook, M., Graham, M., Shelton, T. and S. Gorman (2010). Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake, *World Medical & Health Policy*, 2(2).