*Article*

# Semantic Interoperability of Sensor Data with Volunteered Geographic Information: A Unified Model

**Mohamed Bakillah** [1,2,*]**, Steve H.L. Liang** [1]**, Alexander Zipf** [2] **and Jamal Jokar Arsanjani** [2]

[1] Department of Geomatics Engineering, Calgary University, Calgary, AB T2N 1N4, Canada;
E-Mail: steve.liang@ucalgary.ca

[2] Institute for GI-Science, Rupprecht-Karls-Universität, Berliner Straße 48, D-69120 Heidelberg,
Germany; E-Mails: alexander.zipf@geog.uni-heidelberg.de (A.Z.);
jokar.arsanjani@geog.uni-heidelberg.de (J.J.A.)

**\*** Author to whom correspondence should be addressed; E-Mail: mohamed.bakillah.1@ulaval.ca;
Tel.: +49-6221-54-5563; Fax: +49-6221-54-4529.

**Abstract:** The increasing availability of sensor devices has resulted in important volumes of sensor data, which has raised the issue of making these data fully discoverable and interpretable by applications and end-users. The idea of OGC Sensor Web Enablement (SWE) has addressed this issue by proposing a set of standards to enable accessibility of sensor data over the Web. Similarly, there is a growing interest in volunteered geographic information (VGI). Considering that several researchers have highlighted the potential of this new type of information as a complement to existing, "traditional" data, it becomes important to develop frameworks to support the integration of VGI from several sources and with other types of data. In this paper, we make a first step in this direction by proposing a framework for the semantic interoperability of sensor data and VGI. After having performed an investigation of the types of VGI applications, we have developed a conceptual model of VGI aligned with relevant ISO standards for describing geospatial features. The purpose of this model is to support the generation of common descriptions for VGI applications, which will act as interfaces to higher-level services, such as discovery and reasoning services, in order to be exploited in conjunction with sensor data by client applications. This process is described through architecture for semantic interoperability of sensor data and VGI that we have developed and that we intend to use to set guidelines for future research on integration of VGI in sensor data cyberinfrastructures. We illustrate the

possibilities created by the proposed framework with a description of the various services and interfaces required to implement the framework.

**Keywords:** semantic interoperability; semantic web; sensor data; volunteered geographic information

---

## 1. Introduction

The increasing pervasiveness of sensors and mobile devices has resulted in important volumes of sensor data, which has raised the issue of making these data fully exploitable, *i.e.*, enabling sensor data discovery, sharing, and interpretation. At the same time, there is a growing interest in enabling individuals to voluntarily produce and disseminate geographic information. Such information is currently referred to as volunteered geographic information (VGI) [1]. This information is not produced by individuals or organizations officially tasked to do so, usually designated as data producers, but rather by individuals who operate in various contexts and activities, and who voluntarily gather and share their knowledge on geographic features. For example, the collaborative mapping projects such as Open Street Map (OSM) enable virtually any individual who has Internet access to produce and share maps. As this new VGI paradigm is becoming increasingly popular in geographical information science, several researchers have pointed out the problems raised by this new type of information. For example, some argue that the VGI paradigm is likely to enable the production of larger than ever volumes of geographic information [2,3]. In parallel, VGI has significant differences from traditional geospatial data, which are created by specifically dedicated organizations and experts, according to standardized structures and languages. Indeed, VGI is likely to be produced using natural language rather than formal languages used by existing geospatial databases and services. As a result, some argue that VGI is likely to display higher levels of heterogeneity than traditional data [4]. However, VGI can also be seen as a source of information that can enrich and complete existing geospatial data. For example, VGI can extend existing data sets with information that pertains to temporal and spatial scales and levels of detail that go beyond the capacity of "official" producers [5,6]. In this paper, we also argue that VGI must be fully integrated into existing infrastructures, including infrastructures for sensor data. While numerous approaches to deal with VGI and integrate it into various geospatial applications already exist [3,6–9], there is not a global framework that formalizes the description of VGI applications and their integration into a semantic interoperability process. Despite the fact that VGI applications are obviously accessible through the Internet, there is no framework explaining how different VGI applications can be integrated with sensor data and made accessible through a single portal. Such a framework would enable the full exploitation of VGI and sensor data not only as separated applications and sensor services but as an integrated source of information to support decision-making.

In this paper, we address this issue by proposing a framework for semantic interoperability of sensor data and VGI. While the issue of integrating VGI data with geospatial data in general exists, because of the large scope of this issue, in this paper we chose to focus on sensor data as a first step to explore the integration of VGI with other types of data. Also, VGI displays some similarity with sensor

data, as both can be seen as information that is provided on the environment by "sensing devices" (*i.e.*, either "human sensors" or human-made devices) [1]. Nevertheless, while sensor data can sometimes be considered as VGI, in this paper we will consider them as distinct, because sensor data is not necessarily produced by volunteers and can result from conventional and standardized data productions methods adopted by official sources (governments, mapping agencies), for example, meteorological data. As a result, and in contrast (generally) with VGI, for sensor data it is not necessarily true that the motivation and context of the contributors is unknown. Also, "noise" in sensor data produced by official sources is likely to be lower than that of VGI, even though sensor data can also be affected by quality issues.

This paper makes three main contributions:

- We start by presenting, analyzing and discussing the different types of VGI, especially in order to identify the type of output that they produce. This will set the foundation for developing the model for the description of VGI.
- We have developed a model of VGI that describes the characteristics of these VGI applications and of the data they produce (we use the term "VGI applications" to refer to any software product where humans can use or produce VGI [8]). This model's contribution is to support the management of VGI and its integration into semantic interoperability processes, by providing the conceptual basis for the generation of common descriptions of the heterogeneous VGI applications. These descriptions will act as common interfaces, and their contribution will be to enable the querying and correct interpretation of VGI provided through different applications through a single platform.
- We propose an architecture that explains the process of integrating sensor data and VGI within the same platform and which includes semantic annotation and semantic services to enable the semantic reconciliation of data coming from a variety of sources.

In the last part of the paper, we demonstrate the possibilities created by the proposed framework by describing the various services required to implement the framework with examples. Then, we highlight the research challenges that are yet to be addressed. By doing so, our aim is also to use the proposed framework to set guidelines for future research on integration of VGI in sensor data cyberinfrastructures. In the longer term, our aim is also to highlight the need to develop a standard description model for VGI applications.

This paper is organized as follows: Section 2 provides an overview of the opportunities and challenges brought by emerging VGI applications, as well as a background on geospatial semantics. Section 3 presents the VGI model. Section 4 presents the framework to support semantic interoperability of sensor data and VGI, while Section 5 presents the services required to implement the framework. Conclusions and avenues for future work are provided in Section 6.

## 2. An Overview of VGI: New Opportunities and Challenges

### 2.1. The Emergence of the VGI Paradigm

Back in 1997, Goodchild pointed out that as networks become increasingly ubiquitous, the production of geographic information is moving from a centralized to a distributed process [10].

Nowadays, users can produce geographic information via a variety of Internet applications; as a result, a "global digital commons of geographic knowledge" is created without having to rely solely on "traditional" geospatial data production processes [11]. In 2007, Goodchild introduced the term "volunteered geographic information" to refer to the geographic information generated by users through Web 2.0 era applications. Later, Ballatore and Bertolotto [9] stated that the VGI paradigm reflects the transformation of users from "passive" geospatial information consumers to "active contributors". However, Coleman *et al.* [12] argue that the concept of "user-generated content" is not new, referring for instance to public participation GIS where users can provide input and feedback to decision-makers and involved communities through Web-based applications. The novelty, they claim, lies in part in the community-based aspect of the users' contribution to this digital commons of geographic knowledge [12]: VGI is often created out of the collaborative involvement of large communities of users in a common project—for example, OSM or Wikimapia—where individuals can produce geographic information that emanates from their own local knowledge of a geographic reality or edit information provided by other individuals. Notably, in OSM, users can describe map features—such as roads, water bodies, and points of interest—using "tags", providing information at a level of detail that often goes beyond that which can be provided by traditional geospatial data producers [13].

## 2.2. Opportunities Emerging from VGI

Among the advantages associated with VGI, researchers highlight the use of VGI to enrich, update, or complete existing geospatial data sets [1,5–7,14]. This advantage is especially put forward in the context where traditional geospatial data producers—usually governments—may lack the capacity to generate data sets with comprehensive spatial and temporal coverage and level of detail [5,14]. As a result, Song and Sun [3] indicate that there has been an increase in the usage of VGI in urban management. Furthermore, it was highlighted that VGI can be provided and disseminated in a timely, near real-time fashion, which is required to support decision-making in critical situations such as disaster response and crisis events [1,7,15,16]. In addition, VGI applications allow collecting specific information such as local knowledge, which usually cannot be gathered using traditional data collection processes [1]. The advantages associated with VGI strongly suggest that this type of knowledge is highly valuable and is likely to help provide a dynamic picture of the environment [7,17]. Consequently, researchers start to recognize the need for integrating VGI to existing data sets rather than considering it as parallel information [7,18]. Nevertheless, the integration of VGI into existing geospatial data sets is not yet fully achievable, as it is hampered by various obstacles.

## 2.3. Challenges Related to VGI

Credibility, reliability and quality of VGI are among the main issues being raised [2,19–22]. VGI can be perceived as lacking credibility and reliability because it is produced by non-experts in a context that highly differs from the "structured institution-initiated and expert-driven contexts" [2]. For example, while expert geospatial data producers are expected to generate data with a certain level of precision, users of VGI applications are not formally required to do so, and may have an inaccurate or incomplete perception of the geographic phenomenon they describe. Another concern related to the quality of VGI is the fact that the profile and motivation of contributors are often unknown. As

mentioned by De Longueville *et al.* [7], the socio-economic, sociological and cultural aspects that characterize users can have an impact on VGI generation. Being aware of the relevant characteristics of contributors could help to properly interpret VGI and assess its quality and fitness-for-use. Also related to the issue of VGI quality, the GI community still faces a number of obstacles regarding how VGI can be interpreted, semantically annotated, stored, disseminated, searched, shared and integrated with existing data. Firstly, there are no standard formats and language for collecting VGI and formalizing the context around VGI. This is an initial serious obstacle to the interoperability of VGI with existing data. VGI may be produced and stored using natural language, rather than agreed-upon terminologies and formal language usually employed in existing standardized geospatial database systems [2]. According to Scheider *et al.* [23], the terms used by contributors to describe geographic phenomena lack "unambiguous interpretation in terms of reproducible observations". Nevertheless, some VGI applications, such as OSM, recommend that contributors document and share with others the terms that they use to describe the geographic features, where terms are afterwards organized into a folksonomy. However, Scheider *et al.* [23] indicate that it is difficult to reach a consensus regarding the terminology to use, while Mooney and Corcoran [17] state that there is a lack of mechanisms for checking adherence to the agreed-upon ontology. As a result of this lack of standardization, and also because of the diversity of users' profile and background, some researchers argue that the heterogeneity affecting VGI is likely to be more severe than the heterogeneity affecting traditional geospatial data [4,24]. Therefore, we can expect that semantic interoperability between different VGI data sets can be difficult to establish, since it is already challenging to establish semantic interoperability between official datasets that are based on standardized ontologies or vocabularies. The difficulty to establish semantic interoperability between different VGI data sets affects the ability to control, manage, and distribute the information flow produced by VGI applications, so that it can be fully exploited and used to feed decision-support applications in a timely manner [7]. The issue of semantic interoperability of geospatial data is addressed through the field of geospatial semantics.

### 2.4. Geospatial Semantics and VGI

In the last decade, the semantic web community has been developing theories, techniques and technologies to support semantic interoperability. Semantics is the meaning of symbols in a language [25]. It is fundamental to understand the meaning of data. The main issue that the field of geospatial semantics deals with is semantic heterogeneity. Consider the example of two routing services. One service is based on a data model that considers routes as "practicable by motor vehicles," while the second service considers routes as "practicable by motor vehicles or pedestrians". Running both routing services will not yield the same trajectory because of these different meanings of "routes". The user must be aware of these differences when using the services. The problem of semantic heterogeneity itself is composed of two issues: semantic modeling, and semantic reconciliation [25].

The issue of semantic modeling addresses the problem of which relations and classes should be used to create an abstraction of a geographic phenomenon. This issue is mainly addressed through the research on geospatial ontology engineering [26] and spatial reasoning [27]. Cognitive science demonstrates that semantics is context-dependent and cannot be modeled objectively, nor completely standardized [28]. Therefore, we cannot expect semantic models to be totally compatible. Still,

semantic approaches should be conceived as tools for orienting semantic models so that a basis for comparison exists [28]. With respect to semantic modeling, our contribution in this paper follows this direction, where the model that we propose for the description of VGI applications is intended to provide a common basis for the semantic comparison of different VGI, because semantic heterogeneities cannot be avoided *a priori*. To the best of our knowledge, no such semantic model for describing VGI exists in the literature.

The second issue of semantic reconciliation is that of establishing semantic links between similar pieces of information. It is necessary to enable semantic-based search, integration, and interoperability of geographic information [29]. Semantic reference system is a key concept in the field of semantic reconciliation. Semantic reference systems define a frame of reference where pieces of information from different sources are referenced to a common and formal vocabulary, *i.e.*, a reference ontology that acts as a semantic reference frame (SRF) [30,31]. A semantic reference system includes a function that links a term used in an application (for example, a term used within the description of a sensor) to a concept in the SRF. The linking function can be established with semantic alignment tools. Semantic alignments (also called semantic mappings) are semantic correspondences between elements of semantic models. Semantic alignments allow dealing with semantic heterogeneities between models. Semantic translation, semantic similarity measurement, and geo-ontology alignment are major research fields that deal with semantic alignments [29]. Another research field related to semantic alignments is Linked Data, which its main objective is to provide machine-readable semantic connections between data available on the Web [32,33]. In this paper, the semantic framework that we propose is grounded in key concepts related to semantic reconciliation, including the concepts of semantic alignments, semantic annotations and semantic reference systems. Our contribution is to integrate these concepts into a framework specifically dedicated to the semantic interoperability of VGI, where various VGI sources can be described according to a common model.

To address the challenge of semantic interoperability of VGI, our framework considers the following requirements: it is first needed to make sense of VGI streams [34]. Assigning meaning to VGI streams can be achieved by aligning the semantics of VGI with the geo-ontologies used in existing systems, according to the principle of semantic reference systems [9]. In addition, knowledge extraction methods must be developed to interpret and assign meaning to more complex VGI or VGI coming from multiple sources [7]. Then, a matchmaking mechanism is needed to match VGI with existing data, either for updating or completing these data. Also, VGI needs (expressed by users or systems) must be advertised so that VGI streams can be automatically disseminated to the right systems or users. Addressing the above challenges would allow achieving dynamic semantic interoperability of VGI with existing geospatial data, *i.e.*, ensuring the exchange and integration of VGI from distributed heterogeneous applications with other types of data [35], where systems would be able to "understand" the changes that occur in reality [36]. In this paper, we argue that the lack of a systematic framework that identifies and describes the types of VGI, as indicated by Ostlaender *et al.* [8], is an obstacle to the development of a framework that would support dynamic semantic interoperability of VGI with existing geospatial data. We further argue that such a framework would help to identify and manage the issues pertaining to each category of VGI and their interoperability with existing data. Therefore, this is a motivation for the VGI model that we present in the following.

## 3. VGI Semantic Model

Before presenting the VGI semantic model, we set the grounds by providing an overview of the types of information that are provided through VGI applications.

### 3.1. What Kind of Information Is Provided through VGI Applications?

Today, numerous collaborative Web-based applications allow experts and non-professional users to create and share geographic information [12]. Several VGI applications enable individuals to localize, name and describe geographical features, such as OSM, Wikimapia, and Wayfaring.com. These applications are referred to as collaborative mapping projects. The description of these geographical features can be provided in natural language, or with tags; for example, in OSM, users tag geographical features on the map with key-value pairs (e.g., number of floors: 3 for a building). It is recommended that key-value pairs should be created following OSM's terminology, which is provided through a folksonomy where contributors can add and explain the terms they use [9]. It is estimated that the OSM database contains approximately 1,200,000,000 tagged points, where each point corresponds to a geographic location identified with its latitude and longitude [13]. The data source of OSM is GPS tracks and contributed features by OSM users. The geographical objects can be points and ways. Ways can be arranged to form polygons. The features include roads, points of interest (POIs), transport infrastructure, waterbodies, land use patterns, buildings, *etc*. Wikimapia is another collaborative mapping project. As of today, over 18,000,000 objects have been added in Wikimapia. In contrast with OSM, Wikimapia allows users to delimitate objects with irregular shapes. The Wikimapia API enables acquiring the attributes of the geographical objects such as locations, timestamp, category, tags, and descriptions. Some VGI applications such as OSM 3D also enable their users to specify the 3D geometry of geographical features [6]. Other applications allow users to describe trajectories, e.g., *Endomond,* and *Map My Tracks*, where contributors can share their records of hiking routes, as well as monitor their speed, distance and elevation on their smart phones. Other VGI applications are oriented towards describing more dynamic phenomena; for example, Inrix is an application that collects and compiles the trajectories of vehicles to provide real-time information on evolution of the road network traffic [1]. Early warnings on environmental conditions can also be communicated through VGI applications using various means (voice messages, videos, blog posts, geo-tagged pictures, *etc*.) [1], for instance to monitor forest fires [37]. VGI applications also exist to report on land cover. One example is *Geo-Wiki* (http://geo-wiki.org), which is maintained by the International Institute for Applied Systems Analysis (IIASA) to improve the quality of land cover maps with the contribution of volunteers. Contributors provide information on species distribution, habitat, ecosystems, *etc*. The features created by contributors are polygons representing land cover type.

Weather mapping projects are another example of VGI application, where users provide the current weather conditions in given places. An example of a weather mapping project is 360.org, which gathers weather data from weather stations, universities and amateurs. This application generates geo-located weather reports where the location of the weather station is combined with the volunteered contributions. The resulting type of data is points associated with weather type, temperature, reporting timestamp. Photo sharing websites are also becoming more and more popular. Flickr is an example of

a VGI application where users can share photographs of geographical features geo-referenced by latitude and longitude [1]. As of today, Flickr has more than 6 billion uploaded photos, of which more than 200 million are geotagged. Users can edit photos by adding tags and other textual descriptions. Flickr has been used for various applications including urban characterization [38], pedestrian navigation [39], and event detection [40]. Panoramio is another photo sharing website, which contains geo-located photos. The photos are accessible as a layer in Google Earth and Google Maps and updated regularly. It also provides tags associated with the photos, which can be downloaded through its API. Applications that gather the attitudes and behaviors of users (e.g., regarding use of public transportations) are also considered as VGI applications [41]. VGI applications also exist to monitor various types of sociological phenomena. For example, WikiCrimes provides interactive maps to anonymously report and locate crimes. The geo-located reports are recorded as geographic points. In addition to the location, the type and density of crime can be visualized. WikiCrimes was useful to support research in the field of crime management, e.g., [42]. Finally, VGI applications also enable users to provide quality assessment, for instance, the rating by users of a touristic place [21].

In order to support the development of the VGI Semantic Model presented in the next sub-section, we have classified the types of VGI according to the type of data output that they produce, either VGI provided by sensing devices, geo-referenced text, or geo-referenced features (Table 1). This classification is used in the VGI Semantic Model as a basis to establish the classes of VGI types. Of note is that a lot of VGI applications provide several types of VGI, for example, Flickr contains photos and geo-referenced text.

**Table 1.** Types of volunteered geographic information (VGI).

| Types of VGI | | Examples |
|---|---|---|
| **Type** | **Sub-Types** | |
| VGI provided by sensing device | Pictures | Photo sharing websites (e.g., Flickr, Panoramio, 360cities, MySpace) |
| | Video | YouTube |
| | Sound record | Bird songs (e.g., xeno-canto) |
| | Data stream | Sensor data provided by volunteers (e.g., through open platforms such as www.geocens.ca or 52North), weather reports (e.g., 360.org) |
| Geo-referenced text | | Text ratings of touristic places (e.g., in MouthShut.com, Foursquare), text description of geo-located phenomena (e.g., WikiCrimes) |
| Geo-referenced feature | | Collaborative mapping projects (e.g., OpenStreetMap, Wikimapia, WayFaring.com, GoogleMapMaker), GPS vehicle traces (trajectories, e.g., Inrix), trajectories of outdoor activities (e.g., Endomondo, Map My Tracks), polygons representing land cover type (e.g., GeoWiki) |

*3.2. Description of the VGI Semantic Model*

This section presents the VGI Semantic Model that we have developed in support of semantic interoperability of VGI with sensor data. The model is intended to be used as a common semantic reference frame where VGI provided through any application can be referenced, according to the principle of semantic reference systems. By doing so, VGI can be more easily assigned a formal

meaning through semantic annotation, and then be processed, managed, reasoned with and integrated with sensor data.

Figure 1 below presents the proposed VGI Semantic Model, which is here formalized with UML. The model describes the types of VGI that are produced by VGI applications so that they can be aligned with existing standards. The Figure therefore shows how various classes of the VGI Semantic Model align with the relevant classes of the ISO 19100 series of standards, including the ISO 19109 General Feature Model (GMF), the ISO 19107 Spatial Model and the ISO 19111 for Spatial Reference Systems. This alignment ensures that the proposed VGI Semantic Model is interoperable with protocols used in standard-compliant data infrastructures. We have used the prefixes GF for ISO 19109 classes, GM for ISO 19107 classes, and SC from ISO 19111 classes. The package entitled ISO Standards Feature Model shown in Figure 1 is partly drawn from the framework of Lemmens [43] for interoperability of geo-services. This package is mainly developed around the concept of "feature," which is defined in ISO 19101 as "an abstraction of real world phenomena," and includes geographical entities such as building, water bodies, streets, *etc*. We have distinguished between object features, and event features. Contrarily to the object feature, which is a physical entity located in space and time, the event feature is an occurrent, *i.e.*, a process, event, activity or change that unfolds itself through a period of time [44]. The purpose of this distinction is to allow the separate processing of events that could be inferred from VGI stream, and that could be interpreted, for instance, as alerts or warnings for end-users. The feature is defined with different types of properties, including behavior, association roles with other features and attributes. Association roles are relations between classes (e.g., "close to"), while attributes are characteristics of classes which values are taken from a predefined data type (e.g., GML geometries, date-time, integer, string, *etc*.). The geometric object (including GM_point, GM_line, GM_polygon and GM_volume for 3D objects) instantiates the spatial attribute; the temporal attribute is similarly associated with the relevant classes of the Temporal Model ISO specifications, but for the sake of simplicity, it was not illustrated in this Figure.

The main contribution of the model is included in the package entitled VGI Model, which its intention is to represent any type of VGI. To distinguish the classes included in this package, we used the prefix "VI." The model is developed around the main class named VI_VGI Type. Considering the types of VGI applications reviewed for the purpose of this research and summarized in Section 3.1, we have categorized VGI types into three main categories: VI_Georeferenced Feature, VI_Text, and VI_Sensing Device Output. VI_Georeferenced Feature is a class that represents any geographical feature identified and localized by a user (who instantiates the class named VI_Contributor User) through any application. Because it is geo-referenced, this feature is associated with an instance VI_Location. The VI_Location class is to specify the location of a geo-referenced feature not only with coordinates but also with the name of a place. For example, in MouthShut.com, a website for rating touristic places, text descriptions are associated with place names, but not necessarily with coordinates. Because some applications also allow their users to specify the geometry of identified features, we have created the class VI_Geometric Feature, which is a subclass of VI_Georeferenced Feature. In addition to being associated with a location, an instance of the VI_Geometric Feature class is associated with an instance of the class VI_Geometric Object. Also, in the event of a VGI application that would allow its user to report on moving features (such as vehicles), we have included a class named VI_Mobile Feature, which is associated to the class VI_Trajectory, representing the path

followed by the moving feature in space. The class VI_Trajectory class can be used for example to describe GPS vehicle traces, or trajectories for outdoor activities such as in Endomondo and Map My Tracks. Each instance of VI_Georeferenced Feature can be linked to one or more instances of the VI_Tag class, where a tag is the combination of a label (e.g., type of building) and a value for this label (e.g., hospital). The class VI_Text can be used to represent natural language descriptions associated with geographical locations, such as blog posts or text messages related to a given location. Finally, the class VI_Sensing Device Output includes any VGI provided through a sensing device, including picture, video, sound record, and data stream produced by sensors (which could be wearable sensors).

Of note is that, in the proposed model, we have made a distinction between spatiotemporal trajectories (represented with VI_Trajectory class) and data streams produced by sensor devices (represented with VI_Data stream) which are series of couples with spatial and temporal coordinates in addition to measurements and measurement units. The sensing device could be identified as static or mobile, and therefore, linked respectively to a location or a trajectory. Each instance of the class VI_VGI Type is associated with a context, which is composed of one or more context elements. The purpose of these context elements is to specify any type of information useful to understand the context in which the VGI was created (e.g., the application domain, the intended use of the provided VGI, *etc*.). Each instance of the class VI_VGI Type is also associated with a set of quality elements, which are instances of the class VI_Quality Element. Quality elements could include the accuracy of spatial positioning, vagueness of the geographical area being reported on, *etc*. At last, the VGI application through which an instance of the VI_VGI Type class was provided (e.g., OSM) is explicitly identified, as well as the profile of the contributor, which can include its motivation, expertise, *etc*. Table 2 summarizes the classes of the model and their meaning.

Some classes of the VGI Semantic Model are linked to classes of the ISO Standards Feature Model with a dotted link. This link represents annotation of a class to a standard class and inheritance. Therefore, it represents a subsumption (sub-class) relation. For example, because the class VI_Geometric Feature is linked to GM_Geometric Object, it means that a geometry created through a VGI application is identified as an instance of the class GM_Geometric Object, therefore inheriting its properties, including having a position identified by a set of coordinates in a given spatial reference system. Similarly, the label and label values that compose a tag are linked to the class GF_Feature Type, which indicates that their semantics can be defined with properties including behavior, attributes and association roles. Therefore, rather than only being defined with a term, richer semantics will enable a more accurate interpretation of the meaning of VGI.

In the following, we will briefly discuss how VGI can be referenced to the VGI Semantic Model, followed by the framework that details how VGI is integrated into the semantic interoperability process for sensor data.
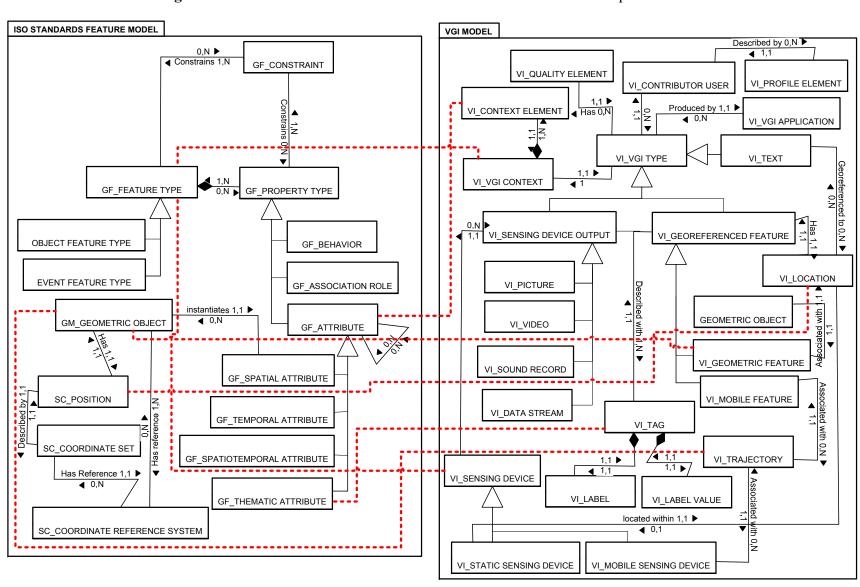
**Figure 1.** VGI Semantic Model and its relation with ISO 19100 series specifications.

**Table 2.** Meaning of the classes of the VGI Semantic Model.

| Class | Meaning |
| --- | --- |
| VI_Context Element | A type of attribute that describes context (task, intended purpose, *etc.*) |
| VI_Contributor User | The ID of a contributor (user) |
| VI_Geometric Feature | subclass of VI_Georeferenced Feature that has a spatial extent |
| VI_Georeferenced Feature | Instantiated by any geo-located features |
| VI_Label | A string used in a Tag as the name of an attribute |
| VI_Label Value | A value for the Label used in a Tag (string or number) |
| VI_Location | The position of a Georeferenced Feature |
| VI_Mobile Feature | A georeferenced feature which position changes over time |
| VI_Mobile Sensing Device | A Sensor Device which position changes over time |
| VI_Profile Element | Attributes describing the profile of the Contributor User |
| VI_Quality Element | An attribute describing the quality of an instance of VGI Type (e.g., accuracy) |
| VI_Sensing Device | A device that records observations of a specific attribute |
| VI_Sensing Device Output | The observation generated by a Sensing Device |
| VI_Static Sensing Device | A Sensor Device which position does not change over time |
| VI_Tag | A combination of Label and LabelValue for describing a Georeferenced Feature |
| VI_Trajectory | The path followed by the Moving Feature in space |
| VI_VGI Application | The name of the source (e.g., OSM) |
| VI_VGI Context | Composed of a set of context elements |
| VI_VGI Type | A root class that encompasses all categories of VGI |

## 4. Toward a Unified Framework for Semantic Interoperability of Sensor Data with VGI

In this section, we present the architecture for realizing the semantic interoperability of sensor data with VGI. This architecture is given in Figure 2. The framework for this architecture is based on existing semantic Web concepts and technologies, including components of semantic reference systems, ontologies, annotations. The main contribution of the framework is to enshrine the VGI Semantic Model within these concepts so that it is eventually compatible with existing technologies.

At the lower level of the architecture, sensors networks and VGI applications enable the collection of various types of data on environmental and urban phenomena. The idea of the architecture is that just as individual sensors and sensor networks must be integrated into sensor observation services (SOS) to be made available on the Sensor Web, and then further be exploited to answer complex queries posed by users, VGI applications must be also made readily accessible and exploitable in a similar manner through standardized interfaces that could be used to describe any VGI application. Then, sensor data and descriptions as well as VGI and VGI application descriptions must be described in terms of a homogeneous terminology in order to resolve semantic heterogeneities. This is done in the semantic annotation layer. As a result, both sensor data and VGI are provided to users through the same portal. The semantic layer includes the more complex semantic reasoning and discovery services that will process sensor data and VGI to answer more complex queries posed by users and decision-makers.
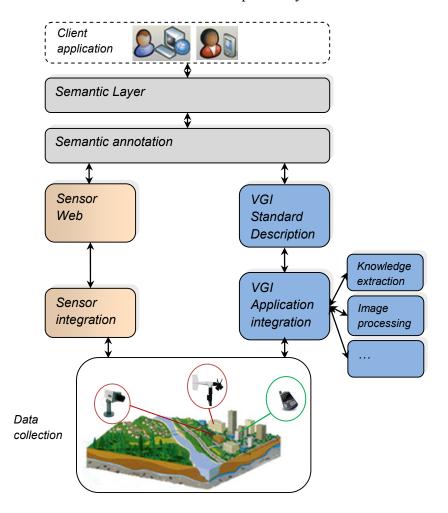
**Figure 2.** Framework for semantic interoperability of sensor data with VGI.



*4.1. Integrating Sensors with the Sensor Web*

Although the objective in this paper is not to focus on the integration of sensors with the Sensor Web or to propose a solution to this issue, in this section we briefly describe this process by referring to state-of-the-art representative approaches. The OGC SWE initiative has developed a set of standards to achieve the so-called vision of the Sensor Web. The latter is aiming to enable the discovery and access of sensor data over the Web [45]. For example, the Sensor Observation Service (SOS) is a standard interface for accessing descriptions of sensors and their observations [46]. While the SOS was initially designed to encapsulate raw data streams from sensors and sensor networks, it has been increasingly employed to give access to more complex information, including processed sensor observations such as aggregated data [7]. The Sensor Event Service (SES) is another example of a proposed SWE standard which allows users to subscribe to events of interest through constraints; the SES monitors the observations produced by registered sensors and notifies users in the same way as a publish-subscribe system would do (OGC 2008) [47]. The Sensor Web is intended to make sensor data interoperable and hide from users and applications the heterogeneities of sensor protocols [45]. To integrate sensors into the Sensor Web, Bröring *et al*. [45] indicate that firstly, we must enable new sensors to automatically register with SWE services that have advertised interest in the sensors' characteristics, such as the type of sensor, the property observed, the location of the sensor, *etc*. Also

needed is a semantic mediator service that performs matchmaking between the sensors' characteristics and the SWE services requested characteristics. Bröring *et al.* [45] also indicate that to integrate sensors into the Sensor Web, we need to enable the exchange of messages between sensors and SWE services; to do so, we must translate between heterogeneous sensor protocols and the SWE protocols by associating with the sensor a generic and sharable interface description. Examples of translation between sensor protocols and SWE protocols can be found in Walter and Nash [45,48,49].

## *4.2. VGI Application Integration*

Similarly to sensors, which must be integrated with the Sensor Web to be accessible, VGI applications must be made accessible by generating common interfaces which will allow the higher-level services to recognize them and understand the kind of data that they can provide. The purpose of the VGI application integration layer is to generate such common descriptions of VGI applications in order to make VGI reusable by these services.

### 4.2.1. Registration of VGI Applications

In order to enable the processing and visibility of VGI, each VGI application should be registered according to a common format. To support this registration, we propose a registering service where VGI application providers can connect and provide the requested characteristics on their VGI application. The VGI application description complies with a profile which is a subset of the VGI Semantic Model. This compliance ensures that the other components of the architecture can access and reuse the characteristics of the registered VGI applications. The required characteristics include the following:

- URL of the VGI application, which will constitute the application's unique identifier (attribute of the VI_VGI Application class)
- Contact information of the company or the person that operates the VGI application (attribute of the VI_VGI Application class)
- Context elements on the VGI application (instances of the VI_Context Element class), including keywords describing:
  - o the intended use(s) of the VGI being provided
  - o the application domain
  - o the geographical area being covered, and
  - o the date when the application was created
- Type of VGI being collected (*i.e.*, sensor device output, geo-referenced features, and/or text)
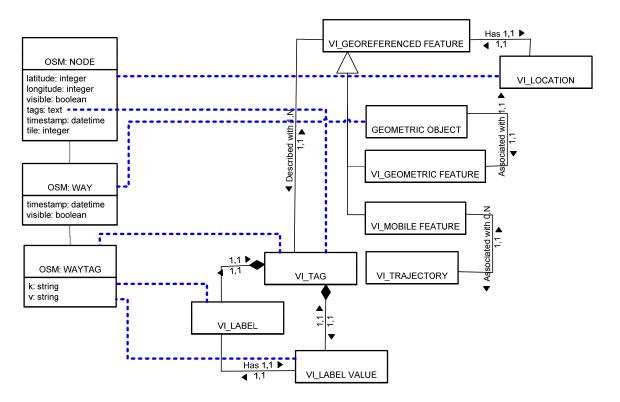
These characteristics represent common metadata which is useful to support the discovery of relevant VGI application. Nevertheless, more specific information on the type of data provided by the VGI application is required to enable retrieval of relevant VGI. This information is contained in the data model of the VGI application.

### 4.2.2. Grounding VGI into the VGI Semantic Model

Once the VGI application is registered, the data model of the VGI application must be linked to the relevant classes of the VGI Semantic Model through mappings. A mapping is a type of relation

between a source element, *i.e.*, an element (attribute, class or relation) of the VGI application's data model, and a target element, which is an element of the VGI Semantic Model. The relation indicates that the source element is a type of target element ("is-a" relation). As an example, Figure 3 shows an excerpt of the OSM data model, where a node corresponds to a distinct geographic location with distinct latitude and longitude values, a way is a combination of points forming either a closed area (*i.e.*, a polygon) or a line. The way tag is the thematic information attached to the way by the contributor, in the format of a key-value pair, where both the key and the value are string data types. The mappings ("is-a" relations) between the OSM data model and the VGI Semantic Model are represented with blue lines. For instance, the OSM node class is a type of VI_Location, while the OSM way tag class is a type of VI_Tag. Besides relations between classes, mixed relations, such as relations between attributes and classes, are also possible. For example, the attribute k (for key) of the class way tag is linked to the class VI_Label. Also, the attribute "tag" of the class OSM node, which is a key-value pair, is linked with VI_Tag.
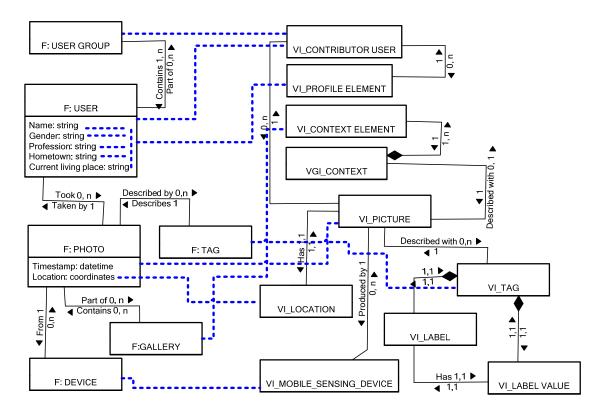
**Figure 3.** Example of mapping between a VGI application model (OpenStreetMap) and the VGI Semantic Model.



We have chosen to illustrate the use of the VGI Semantic Model with OSM because OSM is one of the most used and well-known VGI sources. Another popular VGI example is Flickr. In Figure 4, we have illustrated the use of the VGI Semantic Model for this VGI source, to show that the semantic model is also appropriate for VGI applications other than collaborative mapping projects. The classes of the Flickr data model are distinguished with the "F" prefix. In Flickr, photos are associated with a user, which is described by attributes such as name, gender, profession, *etc*. The user attributes are mapped as VI_Profile Elements of the VI_Contributor User. Users in Flickr can also be part of users groups, where members share common interests. Users groups in the VGI Semantic Model can be

considered as contributor users with VI_Profile Elements being the common interests. This enables to search for VGI (photos) that have been contributed by a particular type of user group. Photos in Flickr can also be part of a gallery, which is a collection of photos related to the same topic. The gallery is here considered as a VI_Context Element that characterizes the photo. Flickr also identifies the device with which the photo was taken, which corresponds to a mobile sensing device.

**Figure 4.** Example of mapping between a VGI application model (Flickr) and the VGI Semantic Model.



To support the establishment of mappings between VGI application data models and the VGI Semantic Model, we propose a mapping registration service where the VGI application provider can visually match the elements of the data model to the elements of the common VGI Semantic Model. Examples of such mapping tools, such as PROMPT, can easily be adapted and deployed to implement this service. In addition, matching algorithms can be deployed to support the establishment of the mappings, such as proposed by Bakillah and Mostafavi [50] or Janowicz *et al.* [51]. To implement the mapping registration service, the VGI Semantic Model was formalized with Web Ontology Language (OWL) and the mappings can be expressed as OWL properties.
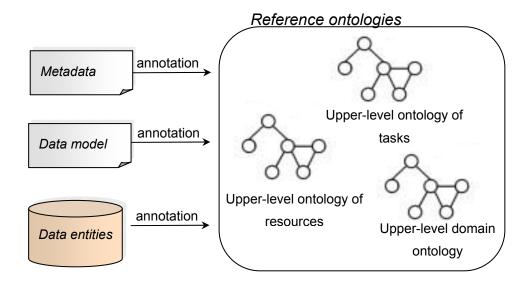
*4.3. Semantic Annotation*

Because each VGI application is made available by different providers for different purposes and in different contexts, their data and metadata models are semantically heterogeneous. For similar reasons, sensor data and metadata models are also semantically heterogeneous [52]. To resolve those heterogeneities, the data and metadata model elements must be referenced to an appropriate common and formal vocabulary, *i.e.*, a semantic reference frame, or reference ontology [53]. The process of

referencing a data model element to a reference ontology is called semantic annotation. Semantic annotation is defined by Klien [54] as the process of establishing explicit relationships (mappings) between elements of the data model to elements of the ontology. Because semantic annotation does not modify the data model, it preserves the semantic independence of application-specific data models [55]. As proposed by Maué *et al*. [56], semantic annotations are established at three different levels: data entities, data models and metadata (Figure 5). Semantic annotations can be stored in different ways. Based on how the semantic annotations are stored, they are referred to using different expressions [43]:

- *Semantic markup*: this expression is employed when the semantic annotations are included within the information source;
- *Registration*: this expression is used when the semantic annotation is stored within the ontology;
- *Registration mapping*: this expression is employed when semantic annotations are stored in a separate source, which contains pairs of identifiers from the information source and the reference ontology.

**Figure 5.** Framework for semantic annotation (adapted from [56]).



In this framework, we propose to employ registration mappings, because semantic markup and registration would involve having a certain control over the information sources and the reference ontologies being used. Therefore, a separate registration mapping store is maintained. Because registration mappings will be employed in various semantic interoperability processes deployed at the semantic layer level (see Section 4.4), registration mappings are expressed with the OWL, a W3C-recommended language for the Semantic Web that supports reasoning. The establishment of semantic annotations can be a very complex process because of the large variety of semantic heterogeneity issues. While the semantic annotations can be established manually with the help of an ontology mapping editor, automatic semantic annotation methods and tools are also desirable because of the large volume of elements to annotate. While it is outside of the scope of this paper to provide a semantic annotation method, we argue that as for semantic mapping, the semantic annotation process should include several techniques to cope with the complexity of the task and the nature of VGI.

Registration mappings at the data and metadata levels involves matching already formalized components (classes, attributes, relations) to concepts of the semantic reference frame. At these levels, semantic annotations can be established using semantic matching tools. Semantic similarity tools can also be deployed to identify the semantic correspondences. Semantic similarity is a quantitative measurement of the descriptions of concepts, which is intended to reflect as much as possible the closeness in meaning. A comprehensive literature review of semantic similarity measures is provided by Schwering [57]. Also, [51] is an example of semantic similarity tool that can compare concepts expressed with Description Logics (DL). Some semantic similarity measures were also developed specifically for VGI. For example, [58] propose a semantic similarity measure for comparing OSM geographic classes. [59] also propose a semantic similarity measure for OSM features that takes into account the history of changes in the naming of these features. At the data entity level (Figure 5), complex knowledge extraction tools are needed to identify entities and the concepts that represent them, for example, to identify geographical features on a picture, to identify movement on a video, to parse text descriptions, *etc*. For example, Klien [54] proposed a rule-based strategy for semantic annotation of data, where the conditions for an object to be an instance of a concept are expressed as a rule (e.g., a river is a waterway whose width is larger than 10 meters). Once semantic annotations are established, the semantics of VGI applications and sensor data sets is readily usable by reasoning systems of the semantic layer.

### 4.4. Semantic Layer

Figure 6 illustrates the components within the semantic layer of the framework architecture.

The semantic layer is designed as a publish-subscribe system, where, on the one hand, VGI providers and sensor services can connect and advertise their capabilities, and, on the other hand, information requestors can advertise their information needs through semantic queries. The information broker is responsible for matching these information needs with the available capabilities.

The formal description of the capabilities of sensor observation services (SOS) corresponds to the output of the SOS, and metadata as provided by SensorML descriptions, whereas the description of the capabilities of VGI providers corresponds to the VGI application profile described in Section 4.2.1. Because the elements provided in these descriptions are semantically annotated, as described in Section 4.3, they are readily interpretable by the information broker to support various reasoning tasks. Semantic queries posed by users can include constraints on various aspects, including:

- Context elements, such as the intended use(s) of the data, the application domain, the geographical area being covered, *etc*.
- Type of data (*i.e.*, videos, pictures, geometries, *etc*.)
- The quality of data
- The entities (objects of events) of interest
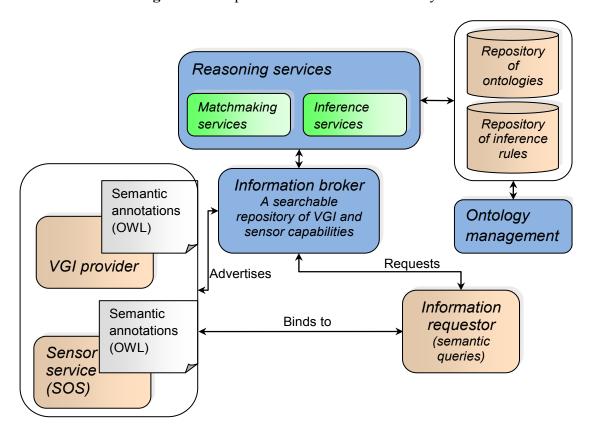- The properties of entities of interest

**Figure 6.** Components within the semantic layer.



The information broker requests reasoning services to determine matches between available data and semantic queries. On the one hand, matchmaking services take as inputs the semantic queries and the capability descriptions, and use a combination of matching algorithms to compute a semantic relation between the semantic query and the capability descriptions. As explained in previous work [50], to be efficient and have a wide scope, the matchmaker component should include different levels of matching, including syntactic, lexical, structural, and semantic matching. It is also desirable that the matchmaking mechanism produces both qualitative and quantitative results (*i.e.*, a semantic relation and a semantic similarity), such as in Bakillah *et al*. [60], to improve the interpretation of the relation between a query and a capability description. On the other hand, the reasoning services include an inference service that uses rules to infer new facts from existing facts. To be compliant with the OWL, the inference service uses the Semantic Web Rule Language (SWRL), which expresses Horn-like rules in terms of OWL classes. SWRL rules can also be employed to express event-condition-action (ECA) rules [61], which in turn can be employed to infer alerts and warnings as requested by users. Reasoning services refer to different ontologies to support inference of new facts, as per knowledge pertaining to various application fields. The ontologies include high-level ontologies with several domains, such as the Semantic Web for Earth and Environmental Terminology (SWEET) ontology (http://sweet.jpl.nasa.gov/ontology/), domain independent lexicons such as WordNet (http://wordnet.princeton.edu/), and application ontologies that describe specific domains, such as ontologies for tasks in disaster management [62]. Ontologies can be edited and managed through ontology management services.

## 5. Application of the Framework

This section describes the services required to implement the framework for semantic interoperability of sensor data and VGI. The required services are listed in Table 2, where the main offered functionalities of these services are provided alongside. The combination of these services generates a semantic interoperability platform where VGI applications can be registered and that supports the combination of data from various VGI applications.

The first service is the VGI application registry service. The objective of this service is to enable the registration of new VGI applications on the platform. The registration involves specifying the characteristics of the VGI application, as listed in Section 4.2.1. The registration service provides an interface where VGI application providers can register their application through the Create capability description functionality. The VGI application registry service also includes a component that advertises the capabilities of VGI applications to the information broker (Figure 6). Figure 7 shows an excerpt of the interface for registering VGI applications. The interface allows the VGI application provider to specify the common VGI application description using a template.

The VGI Semantic Model mediator service is responsible for creating mappings between the VGI application data model and the elements of the common VGI Semantic Model. Its main functionalities include the uploading of the local VGI application model and the creation of mappings. The Lookup description functionality enables the VGI application provider to access a description of an element of the VGI Semantic Model to support the creation of appropriate mappings. For example, Figure 8 shows the OWL mappings produced through the VGI Semantic Model mediator service for the relations between Flickr model and VGI Semantic Model, as illustrated in Figure 4.

The purpose of these mappings is also to enable the user to submit a single query that can be processed against several VGI applications at the same time. For example, consider a user who studies VGI phenomena and who wants to investigate the profile of VGI contributors who contributed to a specific city, regardless of the VGI application. This can be done through submitting the query for VI_ProfileElement, which can be processed against all registered VGI applications that have a mapping to VI_ProfileElement.

The Semantic Annotation Service supports the creation of OWL annotations between, on the one side, the data entities, the elements of the data model, and the metadata, associated either with VGI or with sensor services, and, on the other side, the concepts of the appropriate semantic reference frame. This service enables the uploading of a selected reference ontology (given that several reference ontologies are needed to describe different domains), to lookup the description of concepts in the loaded reference ontology, in order to support the creation of registration mappings between the local terminology used in a VGI application or associated with a sensor service, and the concepts of the loaded reference ontology. This service can also call on the Reasoning Services to perform automatic matching tasks and relieve the user from manually creating registration mappings.
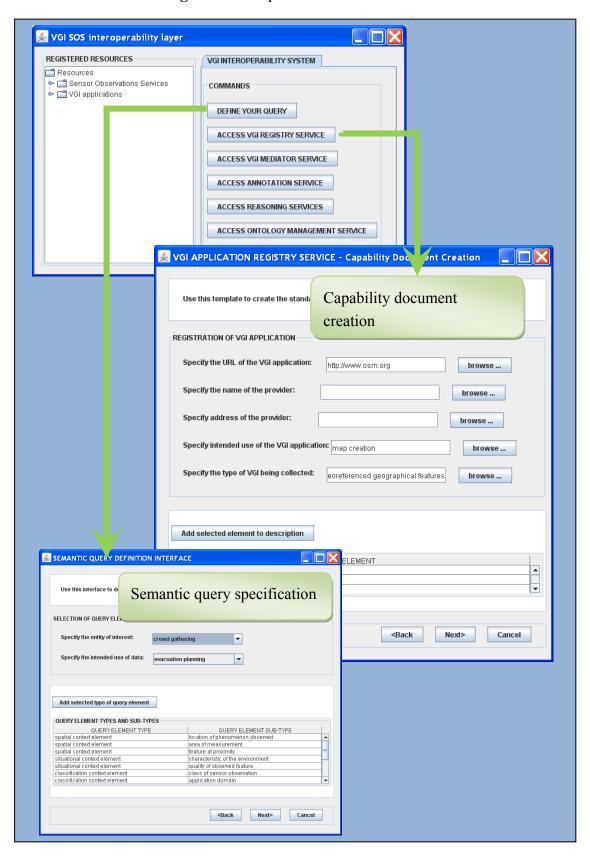
**Figure 7.** Example of service interfaces.

**Figure 8.** Mappings between Flickr and the VGI Semantic Model.

```xml
<owl:Class rdf:about="#FlickrUserGroup">
<owl:equivalentClass rdf:resource="#VI_VIContributorUser"/>
</owl:Class>
<owl:Class rdf:about="#FlickrUser">
<owl:equivalentClass rdf:resource="#VI_VIContributorUser"/>
</owl:Class>
<owl:Class rdf:about="#FlickrUserName">
<owl:equivalentClass rdf:resource="#VI_ProfileElement"/>
</owl:Class>
<owl:Class rdf:about="#FlickrUserGender">
<owl:equivalentClass rdf:resource="#VI_ProfileElement"/>
</owl:Class>
<owl:Class rdf:about="#FlickrUserProfession">
<owl:equivalentClass rdf:resource="#VI_ProfileElement"/>
</owl:Class>
<owl:Class rdf:about="#FlickrUserHometown">
<owl:equivalentClass rdf:resource="#VI_ProfileElement"/>
</owl:Class>
<owl:Class rdf:about="#FlickrUserCurrentLivingPlace">
<owl:equivalentClass rdf:resource="#VI_ProfileElement"/>
</owl:Class>
<owl:Class rdf:about="#FlickrTag">
<owl:equivalentClass rdf:resource="#VI_Tag"/>
</owl:Class>
<owl:Class rdf:about="#FlickrPhoto">
<owl:equivalentClass rdf:resource="#VI_Picture"/>
</owl:Class>
<owl:Class rdf:about="#FlickrPhotoLocation">
<owl:equivalentClass rdf:resource="#VI_Location"/>
</owl:Class>
<owl:Class rdf:about="#FlickrGallery">
<owl:equivalentClass rdf:resource="#VI_ContextElement"/>
</owl:Class>
<owl:Class rdf:about="#FlickrDevice">
<owl:equivalentClass rdf:resource="#VI_MobileSensingDevice"/>
</owl:Class>
```

The objective of Reasoning Services is to provide automatic or semi-automatic matching tools to support the establishment of annotations such as in the above example. It can also include higher level inference processes such as rule-based reasoning. Reasoning Services support uploading a reference ontology, which contains rules to enable inference of facts based on data retrieved from registered VGI applications or sensors services. When both types of data are annotated to the same reference ontology, it is possible to reason with and to combine different data. For example, consider the case of merging data from Flickr with data from OSM, for a scenario where a user wants to find and explore natural reserves in a region. With OSM, he or she can retrieve these regions on the map, but cannot have an idea of how they look at the same time. Semantic annotations to a common reference ontology can support the combination of Flickr data with OSM. Figure 9 is an example of annotation at the metadata model level for OSM. Registration mappings are expressed with OWL class axioms rdfs:subClassOf (to indicate subsumption), or owl:equivalentClass. In this example, the class axioms are applied between OSM features (http://wiki.openstreetmap.org/wiki/Map_Features) and classes of the Semantic Web for Earth and Environmental Terminology (SWEET) ontology (http://sweet.jpl.nasa.gov/

ontology/), which plays the role of the reference ontology for geographic features. The OSM feature "nature reserve" is linked to the LandReserve class of SWEET:

**Figure 9.** Example Registration Mapping from OSM feature class to SWEET ontology class.

```
<owl:Class
rdf:about="http://wiki.openstreetmap.org/wiki/Tag:leisure%3Dnature_reserve">
<owl:equivalentClass
rdf:resource="http://sweet.jpl.nasa.gov/2.2/realmLandProtected.owl#LandReserve"/>
</owl:Class>
```

Figure 10 shows an excerpt of the Realm SWEET ontology, where sub-classes of LandReserve are identified as NationalForest, Park and WildlifeRefuge. By linking OSM features to equivalent and subsumed SWEET classes, OSM objects (e.g., that are instances of OSM "nature reserve" feature) can be linked to entities from other VGI sources that would be identified through semantic annotations to NationalForest, Park or WildlifeRefuge.

**Figure 10.** SWEET ontology LandReserve sub-classes.

```
<owl:Class rdf:about="#NationalForest">
<rdfs:subClassOf rdf:resource="#LandReserve"/>
</owl:Class>
<owl:Class rdf:about="#Park">
<rdfs:subClassOf rdf:resource="#LandReserve"/>
</owl:Class>
<owl:Class rdf:about="#WildlifeRefuge">
<rdfs:subClassOf rdf:resource="#LandReserve"/>
</owl:Class>
```

Consider that Flickr tags and photo descriptions are annotated to SWEET ontologies above, such as in Figure 11 with Forest and Park Flickr tags:

**Figure 11.** Example of registration mapping from OSM feature class to SWEET ontology class.

```
<owl:Class rdf:about="#FlickrTagForest">
<owl:equivalentClass
rdf:resource="http://sweet.jpl.nasa.gov/2.2/realmLandProtected.owl#NationalForest"/
>
</owl:Class>
<owl:Class rdf:about="#FlickrTagPark">
<owl:equivalentClass
rdf:resource="http://sweet.jpl.nasa.gov/2.2/realmLandProtected.owl#Park"/>
</owl:Class>
```

With these annotations, we can easily retrieve the pictures that were taken in the nature reserves identified in OSM and that are about nature reserves or other subclasses of LandReserve in SWEET. These pictures can then be displayed on OSM map, where with only one query the user can retrieve the reserve area on the map, and the corresponding pictures (Figure 12).

**Figure 12.** Combining OSM with Flickr pictures with the support of the Unified Framework.



However, when data from different sources are annotated to different reference ontologies, semantic mappings between references ontologies are needed to enable this type of combination. The reasoning services enable different appropriate matching tools to be called on to support this task. We are currently integrating matching techniques described in previous work on a dynamic semantic mediation service [63]. Also, the OGC Web Processing Service (WPS) standard will be leveraged to implement the reasoning services in order to facilitate the interoperability with other services of the framework. The reasoning services are called by the information broker (Figure 6), which receives the semantic query from the user. An instance of the semantic query interface is also given in Figure 7. The user can select different types of fields (entity of interest, intended use of data, application domain) to further contextualize its query. We also envision that through the semantic query interface, the user could specify thresholds for quality parameters, in order to ensure the fitness-for-use of retrieved data.

At last, the Ontology Management Services support the connection, creation and evolution of reference ontologies. Several ontology editors are already available and could be used to build this service, including Protégé, OntoEdit and SWOOP [64]. In implementing the Ontology Management Services according to the OGC Web Catalog Service standard, we will facilitate the interoperability of the service with other services of the framework. One of the challenges related to the implementation of the framework is also the choice of reference ontologies. It is highly desirable that existing, well-known ontologies be used as reference ontologies to ensure broader facility to establish annotations and mappings. For instance, reference ontologies describing people, their characteristics and relations are necessary to describe users/contributors, for example, the FOAF (Friend of a Friend) ontology, which describes persons, their activities and their relations to other persons or features. Reference ontologies describing geographical phenomena and observations, for example, the Semantic

Web for Earth and Environmental Terminology (SWEET) ontology for observations, to which the categories of features in OSM and tags in Flickr can be linked, are required. Geographic places gazetteer such as GeoNames are also required to link coordinates or place names to common objects.

Many challenges remain to be addressed to fully implement this framework. Among those, the protocols and languages for message passing between services need to be formalized. Further interfaces needs to be developed to support exchanges between VGI applications and the services described in Table 3, for example, to enable the VGI Semantic Model mediation service to upload the VGI application data model. Besides, a lot of research remains to be done regarding the conceptualization of the quality of VGI. In our framework, therefore, further work remains to be done as well to integrate the quality of VGI into the semantic interoperability process. This is a fundamental criterion that will ensure the fitness-for-use of retrieved data and support the use of VGI in decision-making processes.

**Table 3.** Services and functionalities required for the implementation of the framework.

| Services | Functionalities of services |
|---|---|
| **VGI application registry service** | **Create capability description:** enables the provider to describe the profile of the VGI application according to template profile<br>**Advertise capabilities:** the VGI application registry service notifies the information broker of the capabilities of available VGI applications |
| **VGI Semantic Model mediator service** | **Upload local model:** uploads the data model of a VGI application into the VGI Semantic Model mediator<br>**Create mapping:** creates a new mapping between an element of the loaded VGI data model and an element of the VGI Semantic Model<br>**Lookup description:** returns the description of an element of the VGI Semantic Model |
| **Semantic annotation service** | **Upload local terminology:** uploads the terminology (XML) used by the VGI application<br>**Upload reference ontology:** uploads a selected reference ontology from the registry of reference ontologies<br>**Lookup description:** returns the description of an element of the loaded reference ontology<br>**Create registration mapping:** creates a new registration mapping between an element of the loaded terminology and an element of the loaded reference ontology<br>**Call automated semantic annotation tool:** calls a selected automated semantic annotation tool |
|  | **Import output of automated semantic annotation tool:** imports the output relation(s) of an automated semantic annotation tool and inserts the retrieved relation(s) into a registration mapping<br>**Call knowledge extraction tool:** call selected extraction tool<br>**Import output of knowledge extraction tool:** imports the output term(s) of a knowledge extraction tool and inserts the retrieved term(s) into the capability description document |
| **Reasoning services** | **Upload reference ontology:** uploads a selected reference ontology from the registry of reference ontologies into the reasoning service<br>**Select matchmaker:** selects the matchmaker algorithms for a matching task<br>**Call matchmaker:** sends a semantic query and a list of capability description documents to the selected matchmaker<br>**Call inference engine:** sends a set of facts to the inference engine to retrieve new facts |
| **Ontology management services** | **Upload reference ontology:** uploads a selected reference ontology from the registry of reference ontologies into the ontology editor<br>**Create ontology:** creates a new ontology to be added to the registry of reference ontologies<br>**Update reference ontology:** updates the loaded reference ontology |

## 6. Conclusion and Remaining Challenges

In this paper, we have highlighted the need for semantic interoperability of VGI with other types of data. More specifically, we have addressed the issue of semantic interoperability of VGI with sensor data. We have argued, based on existing analysis of the VGI phenomenon, that VGI is a relatively new but useful type of data that brings new opportunities to understand the world in which we evolve, and enable existing data infrastructures with a more "real-world-grounded" representation of the reality. Research on the issue of semantic interoperability of VGI is a first step towards broader acceptance and reuse of VGI in current planning and decision-making activities. Therefore, in this paper, we have proposed a framework for semantic interoperability of VGI with sensor data as a first step into that direction. The framework highlights the need for standards to describe VGI applications as an essential requirement to access VGI from a single portal, integrate VGI with existing data and enable reasoning with VGI. To address this issue, we have proposed a VGI Semantic Model that describes the types of VGI applications and the types of data that can be produced by these applications. We have shown how this model aligns with some OGC standards and therefore, how interoperability can be enabled when VGI applications are described in terms of the elements formalized in this model. The conceptual architecture for semantic interoperability of VGI with sensor data is based on this VGI Semantic Model. It demonstrates how VGI can be integrated along with and in combination with sensor data in high-level reasoning services to support inference of complex facts using Semantic Web technologies.

This framework is a first step towards semantic interoperability of VGI and several challenges remain. This VGI Semantic Model proposed as a basis of the unified framework is intended to be a first step to highlight the need for a standard model for the description of VGI applications. Of note is that since the model has not been extensively discussed with the relevant GI community, it is not intended to be proposed as a standard model. In future work, we propose to expose the model for consultation with relevant stakeholders from the GI community, VGI producers in particular and researchers in the field of VGI. Further requirements for the model can then be identified through the consultation process. Another limitation of the current framework is that we have notably mentioned the need to develop formal communication protocols and languages to support exchanges between the various services that implement the framework. The issue of the quality of VGI has also been mentioned, but extensive work remains to assess, represent, and communicate the quality of VGI. We only mentioned that it is nevertheless a main building block of a successful platform for interoperability of VGI with existing data. Finally, our framework implies the extensive use of various semantic mapping techniques and tools. Extensive research is still needed to make semantic mapping systems truly automatic, as well as to improve precision and recall in less controlled and non-predictable environments. To do so, the challenges that remain to be addressed include, among others: improve the reasoning capability of standard reasoning tools with respect to the spatial and temporal aspects of concepts; integrate more advanced natural language techniques into semantic mapping systems; and develop appropriate external resources (top-level ontologies and thesauruses) to support semantic mapping.

## Acknowledgements

## Conflict of Interest

The authors declare no conflict of interest.

## References and Notes

1. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221.
2. Elwood, S. Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. *GeoJournal* **2008**, *72*, 173–183.
3. Song, W.; Sun, G. The Role of Mobile Volunteered Geographic Information in Urban Management. In Proceedings of the 18th International Conference on Geoinformatics, Beijing, China, 18–20 June 2010; pp. 1–5.
4. Grossner, K.; Glennon, A. Volunteered Geographic Information: Level III of a Digital Earth System. In Proceedings of the Position paper presented at the Workshop on Volunteered Geographic Information, Santa Barbara, CA, USA, 13–14 December 2007; p. 2. Available online: http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Grossner_Glennon_paper.pdf (accessed on 22 March 2012).
5. Tulloch, D. Is volunteered geographic information participation? From vernal pools to videogames. *GeoJournal* **2008**, *72*, 161–171.
6. Goetz, M.; Zipf, A. Towards Defining a Framework for the Automatic Derivation of 3D CityGML Models from Volunteered Geographic Information. In Proceedings of the Joint ISPRS Workshop on 3D City Modelling & Applications and the 6th 3D GeoInfo Conference, Wuhan, China, 26–28 June 2011; pp. 1–16.
7. De Longueville, B.; Annoni, A.; Schade, S.; Ostlaender, N.; Withmore, C. Digital Earth's nervous system for crisis events: Real-time sensor web enablement of volunteered geographic information. *Int. J. Digital Earth* **2010**, *3*, 242–259.
8. Ostlaender, N.; Smith, R.S.; de Longueville, B.; Smits, P. What VGI is Good For-designing a Methodology for Comparative Analysis of Existing Applications to Classify VGI and its Use. In Proceedings of the IGARSS, Honolulu, HI, USA, 25–30 July 2010; pp. 1422–1425.
9. Ballatore, A.; Bertolotto, M. Semantically Enriching VGI in Support of Implicit Feedback Analysis. In Proceedings of the 10th Web and Wireless Geographical Information Systems Conference (W2GIS 2011), Kyoto, Japan, 3–4 March 2011; Tanaka, K., Fröhlich, P., Kim, K.-S., Eds.; LNCS 6574; Springer-Verlag: Berlin/Heidelberg, Germany, 2011; pp. 78–93.
10. Goodchild, M.F. Towards a geography of geographic information in a digital world. *Comput. Environ. Urban Syst.* **1997**, *21*, 377–391.

11. Hardy, D. Volunteered Geographic Information in Wikipedia. Ph.D. Thesis, University of California, Santa Barbara, CA, USA, 2010; p. 260.

12. Coleman, D.; Georgiadou, Y.; Labonté, J. Volunteered geographic information: The nature and motivation of produsers. *Int. J. Spatial Data Infrastruct. Res.* **2009**, *4*, 332–358.

13. Goetz, M.; Lauer, J.; Auer, M. An Algorithm-based Methodology for the Creation of a Regularly Updated Global Online Map Derived from Volunteered Geographic Information. In Proceedings of the Fourth International Conference on Advanced Geographic Information Systems, Applications, and Services, Valencia, Spain, 30 January–4 February 2012; pp. 50–58.

14. Gupta, R. Mapping the Global Energy System Using Wikis, Open Sources, WWW, and Google Earth. In Proceedings of the Position paper presented at the Workshop on Volunteered Geographic Information, Santa Barbara, CA, USA, 13–14 December 2007; p. 2. Available online: http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Gupta_paper.pdf (accessed on 22 March 2012).

15. Palen, L.; Vieweg, S.; Liu, S.B.; Hughes, A.L. Crisis in a networked world: Features of computer-mediated communication in the April 16, 2007, Virginia Tech event. *Soc. Sci. Comput. Rev.* **2009**, *27*, 467–480.

16. De Rubeis, V.; Sbarra, P.; Tosi, P. Web-based Macro-seismic Survey: Fast Information Exchange and Elaboration of Seismic Intensity Effects in Italy. In Proceedings of the 6th International ISCRAM Conference, Gothenburg, Sweden, 10–13 May 2009; p. 11. Available online: http://www.iscram.org/ISCRAM2009/papers/ (accessed on 22 March 2012).

17. Mooney, P.; Corcoran, P. Annotating Spatial Features in OpenStreetMap. In Proceedings of the GISRUK 2011, Portsmouth, UK, 27–29 April 2011; pp. 52–56.

18. Budhathoki, N.; Bruce, B.; Nedovic-Budic, Z. Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal* **2008**, *72*, 149–160.

19. Sieber, R. Geoweb for Social Change. In Proceedings of the Position Paper Presented at the Workshop on Volunteered Geographic Information, Santa Barbara, CA, USA, 13–14 December 2007; p. 3. Available online: http://www.ncgia.ucsb.edu/projects/vgi/docs/supp_docs/Sieber_paper.pdf (accessed on 22 March 2012).

20. Bishr, M.; Mantelas, L. A trust and reputation model for filtering and classifying knowledge about urban growth. *GeoJournal* **2008**, *72*, 229–237.

21. Flanagin, A.; Metzger, M. The credibility of volunteered geographic information. *GeoJournal* **2008**, *72*, 137–148.

22. Gouveia, C.; Fonseca, A. New approaches to environmental monitoring: The use of ICT to explore volunteer geographic information. *GeoJournal* **2008**, *72*, 185–197.

23. Scheider, S.; Keßler, C.; Ortmann, J.; Devaraju, A.; Trame, J.; Kauppinen, T.; Kuhn, W. Semantic Referencing of Geosensor Data and Volunteered Geographic Information. In *Geospatial Semantics and the Semantic Web. Semantic Web and Beyond*; Springer: New York, NY, USA, 2011; pp. 27–59.

24. Hyon, J. Position Paper on "Specialist Meeting on Volunteered Geographic Information." In Proceedings of the Position paper presented at the Workshop on Volunteered Geographic Information, Santa Barbara, CA, USA, 13–14 December 2007; p. 2.

25. Kuhn, W. Geospatial semantics: Why, of what and how? *J. Data Semant. III* **2005**, *3534*, 1–24.

26. Brodaric, B.; Gahegan, M. Ontology use for semantic e-science. *Semant. Web* **2010**, *1*, 149–153.

27. Cohn, A.G.; Hazarika, S.M. Qualitative spatial representation and reasoning: An overview. *Fundam. Informaticae* **2001**, *46*, 1–29.

28. Kuhn, W. Cognitive and Linguistic Ideas in Geographic Information Semantics. In *Cognitive and Linguistic Aspects of Geographic Space, Lecture Notes in Geoinformation and Cartography*; Raubal, M., Mark, D.M., Frank, A.U., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 159–174.

29. Janowicz, K.; Scheider, S.; Pehle, T.; Hart, G. Geospatial semantics and linked spatiotemporal data-past, present, and future. *Semant. Web* **2012**, *3*, 1–13.

30. Kuhn, W.; Raubal, M. Implementing Semantic Reference Systems. In Proceedings of the AGILE 2003—6th AGILE Conference on Geographic Information Science, Lyon, France, 24–26 April 2003; pp. 63–72.

31. Kuhn, W. Semantic reference systems. *Int. J. Geogr. Inf. Sci.* **2003**, *17*, 405–409.

32. Parundekar, R.; Knoblock, C.A.; Ambite, J.L. Aligning Ontologies of Geospatial Linked Data. In Proceedings of the Workshop On Linked Spatiotemporal Data, in Conjunction with the 6th International Conference on Geographic Information Science (GIScience 2010), Zurich, Switzerland, 14 September 2010.

33. Auer, S.; Lehmann, J.; Hellmann, S. LinkedGeoData—Adding a Spatial Dimension to the Web of Data. In Proceedings of the 8th International Semantic Web Conference (ISWC), Chantilly, VA, USA, 25–29 October 2009; Bernstein, A., Karger, D.R., Heath, T., Feigenbaum, L., Maynard, D., Motta, E., Thirunarayan, K., Eds.; Springer Verlag: Berlin/Heidelberg, Germany, 2009; LNCS 5823, pp. 731–746.

34. De Groeve, T.; Stollberg, B.; Vernaccini, L.; Dohorty, B. Mash-up or SDI: Appropriate Mapping Tools for Emergency Situation Rooms. In Proceedings of the 2000 Geomatics for Disaster Management Conference, Torino, Italy, 2–4 February 2010; pp. 122–130.

35. Manso, M.-A.; Wachowicz, M. GIS design: A review of current issues in interoperability. *Geogr. Compass* **2009**, *3*, 1105–1124.

36. Turnitsa, C.; Tolk, A. Battle Management Language: A Triangle with Five Sides. In Proceedings of the Simulation Interoperability Standards Organization (SISO) Spring Simulation Interoperability Workshop (SIW), Huntsville, AL, USA, 2–7 April 2006.

37. De Longueville, B.; Smith, R.S.; Luraschi, G. OMG, from here, I can see the Flames!: A Use Case of Mining Location based Social Networks to Acquire Spatio-temporal Data on Forest Fires. In Proceedings of the 2009 International Workshop on Location-based Social Networks, Seattle, WA, 4–6 November 2009; Zhou, X., Xie, X., Eds.; ACM: New York, USA, 2009; pp. 73–80.

38. Lüscher, P.; Weibel, R. Exploiting empirical knowledge for automatic delineation of city centres from large-scale topographic databases. *Comput. Environ. Urban Syst.* **2012**, *37*, 18–34.

39. Hile, H.; Vedantham, R.; Cuellar, G.; Liu, A.; Gelfand, N.; Grzeszczuk, R.; Borriello, G. Landmark-based Pedestrian Navigation from Collections of Geotagged Photos. In Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia, Umeå, Sweden, 3–5 December 2008; pp. 145–152.

40. Kisilevich, S.; Krstajić, M.; Keim, D.; Andrienko, N.; Andrienko, G. Event-based Analysis of People's Activities and Behavior using Flickr and Panoramio Geotagged Photo Collections. In

Proceedings of 14th International Conference on Information Visualization, London, UK, 26–29 July 2010; pp. 289–296.

41. Birkin, M.; Malleson, N.; Hudson-Smith, A.; Gray, S.; Milton, R. Calibration of a spatial simulation model with volunteered geographical information. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1221–1239.

42. Furtado, V.; Ayres, L.; de Oliveira, M.; Vasconcelos, E.; Caminha, C.; D'Orleans, J.; Belchior, M.; Collective intelligence in law enforcement—The WikiCrimes system. *Inf. Sci.* **2010**, *180*, 4–17.

43. Lemmens, R. Semantic Interoperability of Distributed Geo-services. Ph.D. Thesis, International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands, 2006; p. 323.

44. Grenon, P.; Smith, B. SNAP and SPAN: Towards dynamic spatial ontology. *J. Spatial Cognit. Comput.* **2004**, *4*, 69–104.

45. Bröring, A.; Echterhoff, J.; Jirka, S.; Simonis, I.; Everding, T.; Stasch, C.; Liang, S.; Lemmens, R.; New generation sensor web enablement. *Sensors* **2011**, *11*, 2652–2699.

46. Botts, M.; Percivall, G.; Reed, C.; Davidson, J. OGC Sensor Web Enablement: Overview and High Level Architecture. In Proceedings of the GeoSensor Networks (GSN 2006), LNCS 4540, Boston, USA, 1-3 October 2006; Nittel, S., Labrinidis, A., Stefanidis, A., Eds.; Springer-Verlag: Berlin/Heidelberg, Germany, pp. 175–190.

47. Open Geospatial Consortium (OGC). *OGC Sensor Event Service Interface Specification-Version 0.3.0*; Open Geospatial Consortium: Wayland, MA, USA, 2008.

48. Walter, K.; Nash, E. Coupling Wireless Sensor Networks and the Sensor Observation Service—Bridging the Interoperability Gap. In Proceedings of the 12th AGILE International Conference on Geographic Information Science 2009, Hannover, Germany, 2–5 June 2009; pp. 1–9.

49. Gigan, G.; Atkinson, I. Sensor Abstraction Layer: A Unique Software Interface to Effectively Manage Sensor Networks. In Proceedings of the 3rd International Conference on Intelligent Sensors, Sensor Networks and Information (ISSNIP 2007), Melbourne, QLD, Australia, 3–6 December 2007; pp. 479–484.

50. Bakillah, M.; Mostafavi, M.A. G-map Semantic Mapping Approach to Improve Semantic Interoperability of Distributed Geospatial Web Services. In Proceedings of the Advances in Conceptual Modeling–Applications and Challenges (ER 2010 Workshops), Vancouver, BC, Canada, 1–4 November 2010; Trujillo, J., Dobbie, G., Kangassalo, H., Hartmann, S., Kirchberg, M., Rossi, M., Reinhartz-Berger, I., Zimányi, E., Frasincar, F., Eds.; Springer-Verlag: Berlin/Heidelberg, Germany, 2010; LNCS 6413, pp. 12–22.

51. Janowicz, K.; Keßler, C.; Schwarz, M.; Wilkes, M.; Panov, I.; Espeter, M.; Baeumer, B. Algorithm, Implementation and Application of the SIM-DL Similarity Server. In Proceedings of the Second International Conference on GeoSpatial Semantics (GeoS 2007), Mexico City, Mexico, 29–30 November 2007; pp. 128–145.

52. Bakillah, M.; Liang, S.H.L. Discovering Sensor Services with Social Network Analysis and Expanded Sqwrl Querying. In Proceedings of the 11th Web and Wireless Geographical Information Systems Conference (W2GIS 2012), Naples, Italy, 12–13 April 2012; di Martino, S.,

Peron, A., Tezuka, T., Eds.; Springer-Verlag: Berlin/Heidelberg, Germany, 2012; Lecture Notes in Computer Sciences.

53. Bittner, T.; Donnelly, M.; Winter, S. Ontology and Semantic Interoperability. In *Large Scale 3D Data Integration: Challenges and Opportunities*; Zlatanova, S., Prosperi, D., Eds.; CRC Taylor and Francis: Boca Raton, FL, USA, 2005; pp. 139–160.

54. Klien, E. A rule-based strategy for the semantic annotation of geodata. *Trans. GIS* **2007**, *11*, 437–452.

55. Janowicz, K.; Schade, S.; Bröring, A.; Keßler, C.; Maué, P.; Stasch, C. Semantic Enablement for Spatial Data Infrastructures. *Trans. GIS* **2010**, *14*, 111–129.

56. Maué, P.; Schade, S.; Duchesne, P. *OGC Discussion Paper 08-167r1: Semantic Annotation in OGC Standards*; Technical Report; OGC: Wayland, MA, USA, 2009.

57. Schwering, A. Approaches to semantic similarity measurement for geo-spatial data: A survey. *Trans. GIS* **2008**, *12*, 5–29.

58. Ballatore, A.; Bertolotto, M.; Wilson, D.C. Geographic knowledge extraction and semantic similarity in OpenStreetMap. *Knowl. Inf. Syst.* **2012**, *2012*, 1–21.

59. Mülligann, C.; Janowicz, K.; Ye, M.; Lee, W.C. Analyzing the Spatial-semantic Interaction of Points of Interest in Volunteered Geographic Information. In *Spatial Information Theory*; Egenhofer, M., Giudice, N., Moratz, R., Worboys, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; COSIT 2011, LNCS 6899, pp. 350–370.

60. Bakillah, M.; Mostafavi, M.A.; Brodeur, J.; Bédard, Y. Sim-Net: A view-based semantic similarity model for ad hoc network of geospatial databases. *Trans. GIS* **2009**, *13*, 417–447.

61. Papamarkos, G.; Poulovassilis, A.; Wood, P.T. Event-Condition-Action Rule Languages for the Semantic Web. In Proceedings of the Workshop on Semantic Web and Databases (SWDB 03), Berlin, Germany, 7–8 September 2003; pp. 309–327.

62. Wiegand, N.; Garcia, C. A task-based ontology approach to automate geospatial data retrieval. *Trans. GIS* **2007**, *11*, 355–376.

63. Bakillah, M.; Liang, S.H.L. A dynamic and context-aware semantic mediation service for discovering and fusion of heterogeneous sensor data. *J. Spatial Inf. Sci.* **2013**, *6*, 155-185.

64. Cardoso, J. The Semantic Web vision: Where are we? *IEEE Intell. Syst.* **2007**, *22*, 84–88.