Towards Interoperability: Has theoretical knowledge of Ontologies and Semantics had any impact on Geospatial Applications in GI Science?

Allan Mazimwe and Anthony Gidudu
Department of Geomatics and Land Management, Makerere University

ABSTRACT
The problem faced by Geographic Information Systems (GIS) today is the lack of interoperability among the various systems. Scientists do better when they share resources: computing power, data, tools, models, protocols, and results but making resources available is not the same as making them useful to others. Thus there is need to share common understanding of the structure of information among people or software agents, to enable reuse of domain knowledge, to make domain assumptions explicit and to automatically integrate disparate databases. This research focuses on how theoretical and conceptual research visions in the field of Ontologies and Semantics have impacted on spatial applications today. Using scholar search engines such as Web of Science, Google scholar, Research Gate and GI Science journals, a document review of ontology publications in GI Science was evaluated. Results showed a growing number in Ontology and Semantics publications in the geospatial domain since 1991 and that major research efforts have revolved around creation and management of geo-ontologies, ontology integration, and matching geographic concepts in web pages. Results further showed that ontologies and semantics have been used in SDI implementation, spatial databases, OGC web services, VGI, symbol grounding, semantic similarity, ‘big’ Geodata and sensor networks, location based services, geocoding and so many other applications in the geospatial domain. This shows an evolution in different methods in representing multiple epistemological perspectives of same spatial events and entities as well as attaching contextual information in interest of enhancing interoperability across institutions and geography.

KEYWORDS: GI Science, GIS applications, Interoperability, Ontologies and Semantics.

1.0 INTRODUCTION
Scientists do better science by sharing their resources i.e. computing power, data, tools, models, protocols, and results; but making resources available is not the same as making them useful to others. There is need to share common understanding of the structure of information among people or software agents, to enable reuse of domain knowledge, make domain assumptions explicit and automatically integrate disparate databases. Ontologies have been proposed as a solution to the 'Tower of Babel' problem that threatens the semantic interoperability of information systems constructed independently for the same domain. In information systems research and applications, ontologies are often implemented by formalizing the meanings of words from natural languages (Mark et al., 2003). However, words in different natural languages sometimes subdivide the same domain of reality in terms of different conceptual categories. If the words and their associated concepts in two natural languages, or even in two terminological traditions within the same language, do not have common referents in the real world, an ontology based on word meanings will inherit the 'Tower of Babel' problem from the languages involved, rather than solve it (Mark et al., 2003).

Guarino and Giaretta, (1995) stated that Ontology means something very different in philosophy than it does in information systems. In philosophy, ontology is defined as “what is” while in
Information Science, ontology is defined as “an explicit specification of a conceptualization” (Gruber, 1993), where a conceptualization is a way of “thinking about a domain” (Uschold, 1998) while semantics refers to the meaning of terms. As a GIS community we embrace the information science perspective. The widely accepted conceptualizations of geographic world are fields and objects (Couclelis, 1992 and Goodchild 1992) which are generic conceptual models. Ontologies of the geospatial domain define geographic objects, fields, spatial relations, processes and their categories. Egenhofer and Mark (1995) introduce a body of knowledge that captures the way people reason about geographic space and time. Fonesesca et al (2002) explain the ontology driven GIS architecture that can enable geographic information be integrated in a seamless and flexible way based on semantic values regardless of the representation and for that reason they propose a conceptual model for geographic information with its computer representation. Figure 1 shows the different geographic conceptualizations of same reality and their computer representation stressing the need for ontologies and semantics to ensure interoperability.
According to Schuurman (2006) ontology research in GI Science arguably began in the mid-1990s and three salient issues have been addressed in formal terms through the ontology lens since the mid-1990s namely; Categorization, Data Models, and Semantic Interoperability. And as such there has been a trend in ontology research in that;

- There has been an evolution in different methods of representing multiple epistemological perspectives of same spatial events and entities as well as attaching contextual information to database elements in order to identify different ontologies in interest of enhancing interoperability across institutions and geography (Schuurman, 2009).
- Multiple stakeholders representing different scenarios, agenda and interpretations of the geographical world.

2.0 METHODOLOGY

Using scholar search engines such as web of science, Google scholar, research gate, GI Science journals like International Journal of Geographical Information Science Computers and Geosciences Transactions in GIS, Cartography and Geographic Information Sciences; and international conference proceedings; a count of all publications with the words “GIS/GI Science, ontologies and/or semantics” was made to determine the trends in publications of work related to ontologies and semantics. Furthermore, a search of major top level and domain ontologies in GI Science that have been developed in the last two decades was done to evaluate whether there are researchers who have devoted efforts in the creation of ontologies with a view of explaining the meaning of geospatial concepts. Finally, a document review of publications on applications of ontologies and semantics in GI Science was done together with interviews with GI experts on the applications utilizing ontologies and semantics. The sample of interviewees was randomly selected from GI authors in ontologies and semantics from citation web in the web of science to validate the document review. The applications were then discussed in detail in a view of understanding how theoretical knowledge in the field of ontologies and semantics has had an impact on geospatial applications.

3.0 RESULTS AND DISCUSSION

Results from Web of Science and Google scholar search show a growing trend in ontology and semantics research as shown in figure 2. This is an indicator of growth of theoretical research in the field of Ontologies and Semantics as well as ontology enabled applications.
Search results further indicated that recent GI research has been devoted to developing ontologies with a general view of explaining the meaning of geospatial concepts leading to development of top level ontologies and domain ontologies that are compliant to the W3C standards stack for the semantic web. Such ontologies include;

- SUMO (The Suggested Upper Merged Ontology)
- SWEET (Semantic Web for Earth and Environmental Terminology)
- DOLCE (Descriptive Ontology for Linguistics and Cognitive Engineering) *(Sieber, Wellen and Jin, 2011)*
- DIGEST (Feature and Attribute Coding)
- USGS Spatial Data Transfer Standard (SDTS)
- Geographic Data Description Directory (GDDD)
- Alexandria Digital Library feature Type Thesaurus
- GEMET (General Multilingual Environmental Thesaurus)
- AGROVOC (Agricultural Information Management Standards)
- EuroVoc (Multilingual Thesaurus of the European Union)
- Tired ontologies

This indicates that researchers have considered use of ontologies as a means of knowledge sharing among different geographical databases *(Smith and Mark 1998, Fonseca and Egenhofer 1999)*. Results from the search and interviews further revealed that the GI Research visions arising from ontologies and semantics research in GIScience include; SDI (Geoportals), OGC web services, big Geodata, volunteered Geographic information (VGI), symbol grounding, Digital earth research initiative, sensor networks, spatial databases, semantic web, linked data, and web 2.0, Mobile computing and location-based services, Geocoding, and semantic similarity. Therefore, several semantic conceptual and interoperability frameworks have been proposed in recent years, in the geospatial domain *(Bishr 1998; Brodeur et al., 2003; Rodriguez and Egenhofer 2003; Ahlqvist 2005; Kuhn 2003; Lutz et al. 2003; Kavouras et al., 2005; Lutz and Klein 2006; Bian and Hu 2007; Hess et al., 2007; Cruz and Sunna 2008; Schwering 2008; Staub et al., 2008)*. Below is a discussion of some concepts and applications in GI science that have resulted from ontologies and semantics research and how they have had impacts on several applications in GIScience.

### 3.1 Semantic Similarity
Due to their analogy to spatial proximity functions, semantic similarity measures have been widely studied and applied in GIScience (Rodríguez and Egenhofer, 2004; Li and Fonseca, 2006; Raubal and Adams, 2010). Over the last years, the concept of semantic similarity has gained attention as a non-standard inference service for various kinds of knowledge representations and as such, semantic similarity measurement is a key methodology in various domains ranging from cognitive science to geographic information retrieval on the Web (Janowicz et al., 2008) which is used to measure the degree of potential semantic interoperability between data or different geographic information systems (GIS). The power of similarity lies in providing a graded structure instead of a rigid Boolean matching (Janowicz et al., 2009). Similarity is essential for dealing with vague data queries, vague concepts or natural language and is the basis for semantic information retrieval and integration (Schwering 2008). A number of studies have focused on measuring user similarity using various methods like Feature-based approach (Tversky 77; Rodriguez and Egenhofer 2004), Network approach as modified by Janowicz (2010), geometric approach to trajectory, semantic annotations for venue categorization where the concept of semantic signatures can be applied in assessing user similarities (Ye et al., 2011, Li et al., 2008, Lee et al., 2007), Alignment models and transformation models. The Semantic similarity concept has impacted spatial applications in that it enables efficient:

- Geographic Information retrieval (e.g. in web services and emergence response), theories established for (geographical) information retrieval and in the cognitive sciences use similarity functions to mimic human similarity reasoning as influenced by language, age, and cultural background data integration. Figure 3 below shows two web-based user interfaces implementing similarity and subsumption-based retrieval (Janowicz et al 2009) which have been implemented and are available as operational free and open source software.

The concept of similarity measures is also being used in assessing user similarity for personalized recommender location based systems and Geographic Information retrieval as illustrated by Mckenzi et al., (2013) based on social data such as twitter, foursquare face book etc.

- Ontology merging,
- Searching and browsing through ontologies (semantic web) and hence in knowledge acquisition (Janowicz et al., 2008).
- Spatial decision support (identifying regions with certain characteristics) (Janowicz et al., 2008).
- Volunteered geographic information
- Land mark based navigation illustrated by where facades of buildings locally dissimilar to the neighbouring facades, were selected as prominent landmarks for route instructions in a pedestrian navigation service (Janowicz et al., 2008)
3.2 Semantic Sensors
The use for networking sensors and measurement in GIScience is increasing now and again, this implies an increase in volumes of data thus heterogeneity of devices, data formats and measurement procedures. As such the OGC created Sensor web enablement initiative which caters for syntactical interoperability between sensors. But to achieve full inter-operability and management of the large volumes of data as well as sensor interoperability, there is need for adding semantic interoperability element.

Semantic Web technologies have been proposed as a means to enable interoperability for sensors and sensing systems. Semantic technologies can assist in managing, querying, combining sensors and observation data (Compton et al., 2012). And as such, the Semantic Sensor Network (SSN) incubator group has developed a formal OWL DL ontology for modelling sensor devices (and their capabilities), systems and processes. This semantic sensor network ontology is based on concepts of systems, processes and observations as well as revolves around the central Stimulus-Sensor-Observation pattern. While developing this ontology, a thorough consideration of previous sensor ontologies and concurrent development of informal vocabulary of the main terms, drawing on earlier vocabularies like OIML/VIM and OGC/SWE (SensorML and OandM) was done. Examples of applications where such ontologies have been used include a semantic web based GIS application for environmental management (Tanasescu et al. 2006), and disaster emergency management. Figure 4 shows an application that was developed under sensei FP7 EU project where linked-data location information were associated to local location ontology descriptions enabling navigation through a set of linked sensor data and querying sensors based on their deployment and other attributes as well as their physical locations.
3.3 Big Linked Geo Data

Big data is often characterized by *volume* (size of data, its multidimensional nature and interlinkage in the global graph), *variety* (data formats, social media data, and authoritative data) and *velocity* (speed at which the data is created and updated). With big data, there is a rapidly increasing information universe with new data created at a speed surpassing our capacities to store it, therefore it requires improved methods to retrieve, filter, integrate, and share data (Janowicz 2012).

While the Web has changed with the advent of the Social Web from mostly authoritative towards increasing amounts of user-generated content, it is essentially still about linked documents. In contrast, the upcoming Data Web is about linking data and not about linked documents. With a growth rate of millions of new facts encoded as Resource Description Framework (RDF)-triples per month, the Linked data cloud approach allows users to answer complex queries spanning multiple sources. Due to the uncoupling of data from its original creation context; semantic interoperability, identity resolution, and ontologies are central methodologies to ensure consistency and meaningful results. Prominent geo-related Linked data hubs include Geonames.org as well as the Linked Geo Data project, which provides a RDF serialization of Open Street Map. These hubs have ontologies e.g. Geonames ontology which ensures meaningful and consistent results.

3.4 Geoportals/OGC web Services: Semantic interoperability of web services

---

1 http://www.w3.org/2005/Incubator/ssn/wiki/images/9/96/Mashup-uni-deployment.png
2 http://linkedgeodata.org/About
Although OGC web services have undoubtedly improved the sharing and interoperability of spatial information there are limitations such as difference in semantics in data from different sources (i.e. semantic heterogeneity). As a result, it is difficult for users to automatically compose and perform context-based search thus issues of low recall or low precision. Ontologies and semantics have been identified to overcome this problem of semantic heterogeneity and thus enabling searching, querying and discovery of spatial information. This necessitated research that led to establishment of semantic enablement layer for OGC services. This is achieved by first encoding data and service protocols linked to formal specifications stored in ontologies using annotations; secondly a service has to be established for managing and maintain these ontologies and finally encapsulate the semantic web reasoners to integrate them into SDIs. This research also led to several new services and tools such as conceptVISTA for ontology creation and visualization, the SWING concept repository, to name but a few.

In the Spatial Data Infrastructures (SDI) realm, the research community is aware of the potential benefits of using ontologies as a knowledge representation mechanism (Sen et al., 2007). Ontologies in SDI are used in data sharing and systems development, facilitating resources and information retrieval, and discovery of web services. One typical SDI benefiting from ontology research is the INSPIRE portal which has an ontology-based architecture for Geographic Information (GI) discovery and GI retrieval.

Another application area for ontologies and semantics is the spatial decision support (SDS). Spatial decision support is a dynamic and heterogeneous domain that benefits from a detailed description of its existing process workflows, methods and tools (Li, et al., 2012). Ontologies cover various aspects of spatial decision support ranging from decision problems, processes, methods and technology, over tools, models and data sources, to relevant case studies and literature (Li, et al., 2012). In other words ontologies and semantics support the documentation and retrieval of dynamic knowledge in SDS by offering flexible schemata instead of fixed data structures. And as such SDS ontology was developed. Figure 6 shows architecture of an SDS Knowledge Portal application that incorporates ontologies and semantics to enable search, query and discovery of knowledge.

![Figure 5: System architecture and workflow of the SDS Knowledge Portal application, (Source: Li et al, 2012)](image-url)
An example of such applications was developed under the harmonISA project whose goals was to automatically integrate land-use and land-cover data in the three regions of Friuli Venezia-Giulia (Italy), Slovenia and Carinthia (Austria) as shown in figure 6. The HarmonISA provides several applications and services which include OGC Web Map Services, an ontology viewer, land use ontologies and a harmonized land use viewer a web application with which one can query, filter and navigate through land uses. It addresses the need to add semantics to land use classes in order to enable seamless integration of disparate data sources (in relation to trans-boundary issues)

![Figure 6: HarmonISA land cover viewer. Source: HarmonISA website](3)

### 3.5 Symbol Grounding

There has been much discussion about the scope and limits of purely symbolic models of the mind and about the proper role of connectionism in cognitive modelling. Research has proposed that ontologies are part of the solution to the symbol grounding problem:-How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols? (Hanard, 1990). The concept of symbol grounding in GI Science has been applied in semantic image interpretation by Heudelot et al (2004) where there is mapping between numerical image data and high level of semantic representations thus a step towards automated extraction of meanings(semantics) of an image.

### 3.6 Semantic interoperability in spatial databases

---

3 [http://harmonisa.uni-klu.ac.at/harmonisa/application.jsp](http://harmonisa.uni-klu.ac.at/harmonisa/application.jsp)
Geospatial semantics is an emerging research theme in the domain of geographic information systems and spatial databases. Currently, we observe a wide use of geospatial databases that are implemented in many forms. Ontologies play an important role in enabling semantic interoperability between agents by providing them a common understanding of the reality. Ontologies contain some elements of context which are usually defined in the assumptions that help in interpretation of concepts. Ontology of geospatial data cubes would include definitions, assumptions, and properties (spatial and non-spatial) of the data cubes concepts (Sboui et al., 2007). Research in ontology and semantics has been key in providing a conceptual framework and models for supporting semantic interoperability in geospatial data databases/data cubes. Interoperability in these databases helps in enabling simultaneous and rapid navigation through different data cubes, rapid insertion and retrieval of data in a data cube, interactive and rapid analysis of phenomena changes. An example of such conceptual frameworks is provided by Sboui et al., (2007) based on communication patterns between people.

Today, the most successful applications in the ontology field as far as spatial data bases are concerned are database interoperability, and cross database search.

3.7 Semantics of Volunteered Geographic Information

With the advent of Volunteered Geographic Information (VGI) as coined by Goodchild (2007) where volunteers contribute to geographic data and access under public license via platforms like Open-street map, there are a variety of conceptualizations stemming from the variety of data sources e.g. if one wanted to map a certain disaster area, this may be conceptualized very differently by communities thus there is need for some kind of standardization. If one is to map and use VGI, then it is essential that one understands the different conceptualizations and can as well map between them as well as the data, semantics of volunteered geographic information. With increasing success VGI in different domains like disaster management (known as crowdsourcing), and thus being vital in daily lives of citizen, there need to ensure information obtained is quality and useful rather than focus on coverage. Semantic Web technologies have been proposed as a means to enable interoperability for sensors for VGI and systems. According to Kuhn (2007) some of the challenges range as far as exploiting the grounding effect of VGI on semantics; enabling and capturing semiosis in the social networks around VGI; and combining ontologies with folksonomies exist. Folksonomies use user applied tags of document for searches with in the document but the decision to use ontologies or folksonomies depends on application. AVGII/VGS framework ontology has been developed (Savelyev et al 2011) where a linked data model is used. The framework ontology is able to incorporate external vocabularies e.g. FOAF (friend of friend) and because of VGI ontologies there has been rapid urge to develop ontologies tailored to needs of mobile applications coupled with best practices because mobile phones are the major platforms for VGI.

3.8 Geocoding

Geocoding is the act of turning descriptive locational data such as a postal address. Some of the geocoding services include Geonames (open source), yahoo, Google, Open-street map. Previously, the existing geocoding services were generally limited to assign a geographic coordinate to an absolute location such as a street address. Conventional geocoding services that work with absolute locations may not be able to determine the coordinates of the place of incident

---

4 http://www.foaf-project.org/
always. For example, a geocoding service might assign coordinates of a particular place to another
country or place with a similar name which is not right. In order to obtain exact longitude and
latitude of a location, geolocation via ontologies and semantics is necessary. For example one can
search for a place using the Geonames geocoding services, what happens here is that the search
service returns the results as defined by Geonames semantics web ontology. This ensures that
context is attached on the result being given depending on the quality of service.

4.0 CONCLUSION
The problem faced by Geographic Information Systems (GIS) applications today is the lack
interoperability among the various systems. Applications may use different terminologies to
describe the same conceptualization. Even when applications use the same terminology, they often
associate different semantics with the terms. This obstructs information exchange among
applications. The role of semantics for interoperability and integration of heterogeneous data,
including geospatial information, has been long recognized (Sheth 1999). Research in the field of
ontologies and semantics has enabled sharing of common understanding of the structure of
information among people or software agents, to enable reuse of domain knowledge, to make
domain assumptions explicit. Indeed, there is no doubt that that the field of ontologies and
semantics has had an impact on applications in GI Science. This paper highlights the various
research visions in the field of ontology and semantics for GI Science and gives examples how
they impact various Geospatial applications. Major research efforts in ontologies and semantics
have revolved around 1) creation and Management of geo-ontologies that constitutes activities
involved in ontology management including designing, developing, registering, storing,
discovering, visualizing, querying, and maintaining ontologies, 2) Matching geographic concepts
in web pages to geo-ontologies 3)ontology and semantics integration thus new research visions/application areas in GI like OGC web services, SDI,SDSS, geocoding, big Geodata, linked
geo data, volunteered Geographic information, image interpretation and so many other areas of GI
Science. In all these application areas, the major role played by ontologies and semantics is that of
enabling interoperability. In this era of big data, we envisage that future work in will focus
ontology driven data science especially with real time or near real time streams of data, and
geospatial modelling especially on the web with hope of resolving semantic ambiguity.
5.0 REFERENCES


Janowicz, K. 2012, Big Data in Geographic Information Science Panel 2012 In conjunction with the Geographic Vocabularies Camp Santa Barbara 2012 (GeoVoCampSB2012).


Li Quannan, Zheng Yu, Xie Xing, Chen Yukun, Liu Wenyu, and Ma Wei-Ying. 2008. Mining User Similarity Based on Location History. Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems - GIS ’08, page 34,


Schuurman N., 2009. Spatial Ontologies. *Simon Fraser University, Burnaby, BC, Canada*


Smith B, and Mark D 1998. Ontology and Geographic Kinds. In *international Symposium on Spatial Data Handling, Vancouver, Canada*, pp. 308-320


