Verlag Dr. Hut

Khatereh Polous

Event Cartography: A New Perspective in Mapping

GEOWISSENSCHAFTEN

TECHNISCHE UNIVERSITÄT MÜNCHEN

Ingenieurfakultät Bau Geo Umwelt

Lehrstuhl für Kartographie

Event Cartography

A New Perspective in Mapping

M.Sc. Khatereh Polous

Vollständiger Abdruck der von der Ingenieurfakultät Bau Geo Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation

Vorsitzender:

Univ. -Prof. Dr.-Ing. Uwe Stilla

Prüfer der Dissertation:

1. Univ. -Prof. Dr.-Ing. Liqiu Meng
2. Univ. -Prof. Dr. Jukka Matthias Krisp, Universität Augsburg
3. Univ. -Prof. Dr. Alexander Zipf, Ruprecht-Karls-Universität Heidelberg

Die Dissertation wurde am 22.09.2015 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät Bau Geo Umwelt am 27.01.2016 angenommen.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über
http://dnb.d-nb.de abrufbar.

ISBN 978-3-8439-2614-0

Sternstr. 18, 80538 München
Tel.: 089/66060798
www.dr.hut-verlag.de

1. Auflage 2016

ABSTRACT

Depiction of real world within Geo Information System (GIS) presumes the world as some successive and static snapshots rather than a very dynamic system populated by various open-ended phenomena. Due to limitations inherited to the traditional GIS, development of a conceptual model which enables the GIS to record and represent historical states of the globe along with its dynamic aspects is still at its early stages.

In this research, the concept of "mapping" goes beyond the principle of mapping objects which have distinct spatial, temporal and attributive identities as usual in object-oriented systems. The main goal of this thesis is to present a conceptual model for geo-knowledge. This model can handle real world dynamisms such as relationships amongst objects with each other, objects with events, events with events and the involving processes. This study uses a generic event-oriented perspective to implicitly represent causal relationships among different components of a Spatio-Temporal Information System. From this new perspective, the objects in space and time are considered merely as information elements of the events, which are connected to other event elements through internal or external processes.

Making use of the unique opportunity offered by the concept of Volunteered Geographic Information (VGI), a framework was developed to provide the possibility of collecting and storing event related information elements in OpenStreetMap (OSM) platform. These information elements are recorded in OSM history file by adding new key value pairs adequately designed for handling spatio-temporal and semantic information. The implemented framework enables the volunteers and users to map and represent different event elements such as involving objects, event duration, event pattern, and involving participants.

In order to enrich the event database (collected manually by volunteers), a crawling framework was developed to automatically collect freely available event information in various webpages on internet. To utilize the collected event information, a Volunteered Location Based Service (VLBS) – OpenEventMap – was developed. The service provides users the possibility of mining the event database and answering the spatial, temporal and attributive queries associated with the proposed event definition.

The last part of the thesis is devoted to changing the role of the public from pure sensors in the VGI concept. With this regard, the public not only observes environment and collects the information through the senses, but also is able to understand the meaning of the gathered data. For this purpose, an Event Visualization and Analysis Tool (EVT) was developed to increase the general awareness of the volunteers in spatial analysis and shift their power and influences from a mere data collector to a knowledge producer. This accelerates the transition of mastery from professional cartographers to the public; indeed everybody who can understand the space will be able to collect, analyze, evolve and finally synthesize the spatial knowledge. In this way, VGI can move beyond the raw spatial data to realize Volunteered Geographic Understanding.

ZUSAMMENFASSUNG

Die Abbildung der realen Welt in geographischen Informationssystemen (GIS) wird in der Regel als aufeinanderfolgende, statische Schnappschüsse dargestellt, anstatt als eines sehr dynamischen Systems welches verschiedene zeitliche Phänomene beinhaltet. Aufgrund von Einschränkungen der traditionellen geographischen Informationssysteme, wie zum Beispiel der Mangel an Datenverarbeitungsressourcen und den Herausforderungen bei der Entwicklung eines konzeptionellen Modells, in dem ein GIS in der Lage ist sowohl den historischen Zustand der Welt als auch deren dynamischen Aspekt aufzunehmen und darzustellen.

In dieser Forschungsarbeit geht das Konzept des "Mappings" über das Prinzip der Zuordnung eines Objektes als eine konzeptionelle geographische Einheit, die eine eigenständige räumliche, zeitliche und attributive Identität besitzt, hinaus. Traditionelle objektorientierte Systeme beinhalten "low-level" geodatenbasierte Systeme, welche in der Lage sind den Objekten Elemente wie Zeit, Ort und Attribute zuzuordnen. Daher ist eines der Hauptziele dieser Arbeit ein konzeptionelles Modell vorzustellen, welches ein "high-level" geowissensbasiertes System beinhaltet, um die realen Prozesse, wie die Beziehungen zwischen den Objekten untereinander, die Beziehungen zwischen den Objekten und die Beziehungen zwischen den Ereignissen untereinander einschließlich der beteiligten Prozesse einzubinden. Diese Studie verwendet eine generische ereignisorientierte Perspektive, um die kausalen Zusammenhänge zwischen den verschiedenen Komponenten eines räumlich-zeitlichen geographischen Informationssystems darzustellen. Diese neue Perspektive berücksichtigt die raumzeitlichen Objekte lediglich als Informationselemente der Ereignisse, die mit anderen Ereigniselementen durch interne oder externe Prozesse verbunden sind.

Aufgrund der Möglichkeiten, die durch das Konzept "Volunteered Geographic Information" (VGI) entstanden sind, wurde ein Framework entwickelt, was die Möglichkeit zur Erfassung und Speicherung von ereignisbezogenen Informationselementen in der OpenStreetMap-Plattform (OSM) bereitstellt. Diese Informationselemente werden in einer OSM-history-Datei durch das Hinzufügen neuer Schlüssel-Wert-Paare gespeichert. Diese sind entsprechend für den Umgang mit räumlich-zeitlichen und semantischen Informationen gestaltet. Das implementierte Framework befähigt die Nutzer verschiedene Ereigniselemente einzubinden. Dazu gehören zum Beispiel die betroffenen Objekte, die Ereignisdauer, das Ereignismuster und die involvierten Teilnehmer.

Um die von freiwilligen Nutzern manuell befüllte Ereignisdatenbank zu ergänzen, wurde ein "crawling-framework" entwickelt, um frei verfügbare Ereignisinformationen auf ausgewählten Webseiten im Internet automatisch zu sammeln und zu speichern. Um die Ereignisinformationen bereit zu stellen, wurde ein Volunteered Location Based Service (VLBS) namens OpenEventMap entwickelt. Der Service bietet den Nutzern die Möglichkeit, die Ereignisdatenbank zu analysieren und räumliche, zeitliche und attributive Anfragen zu stellen.

Der letzte Teil dieser Forschung befasst sich mit den freiwilligen Nutzern und deren traditionelle Rolle als reine Sensoren in dem VGI-Konzept. Aus einer erweiterten Perspektive sind die freiwilligen Nutzer nicht nur in der Lage, die Umwelt zu beobachten und die Informationen über ihre Sinne zu sammeln, sondern auch die Bedeutung der gesammelten Daten zu verstehen. Zu diesem Zweck wurde ein Ereignisvisualisierungs- und Analyse-Tool (EVT) entwickelt. Dieses Tool zielt darauf ab das allgemeine Bewusstsein der freiwilligen Nutzer für die räumliche Analyse zu sensibilisieren. Damit erweitern die Nutzer ihren Einfluss vom reinen Datensammler zum Wissensproduzenten. Dies

beschleunigt den Übergang der Beherrschung dieses Gebiets von professionellen Kartografen hin zur Öffentlichkeit. Alle die Räumlichkeit verstehen können, sind in der Lage das räumliche Wissen zu sammeln, zu analysieren, zu entwickeln und schließlich zu synthetisieren. Auf diese Weise kann sich VGI über die rohen Geodaten hinaus bewegen und ein Volunteered Geographic Understanding erreichen.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Prof. Dr.-Ing Liqiu Meng whose Big Mind and Heart helped me to fulfill my dream. Your abundant insight, wonderful support and willingness to share your knowledge were invaluable in my research work. I am also very grateful for a very productive and family like working environment you offered us at the department of Cartography.

I am extremely thankful for the mentorship provided by Prof. Dr. Jukka Mathias Krisp, whose immense support was immeasurable. I would like to heartily thank you for always listening to me and giving me words of encouragement even after moving into Augsburg. Your valuable advices, suggestions and comments made a great influence in the completion of this work.

I would also like to thank my sponsor institute "International Graduate School of Science and Engineering, Technische Universität München (IGSSE)". And many thanks to IGSSE staff especially Mrs. Jo-Anna Küster for her great supports and patience during my work at Technical University of Munich.

Special thanks to Dr. Sarah Jane Rehnborg at School of Public affair, the University of Texas at Austin. I learned so much through the opportunity you provided me in Austin. Your expertise in volunteerism and enthusiasm in lifting the other's life opened up a new horizon in my life.

Many thanks to Dr. Peter Mooney at Department of Computer Science, National University of Ireland Maynooth (NUIM). Our research meetings during my stay in Maynooth were invaluable for increasing my insight in the field. Indeed you boosted my research at its very early stage and shed light on my way.

And finally, to my real Love, Amirhani, I am so grateful and thankful for having you in my life journey. This work would have not been possible without the beautiful love and light you always shared with me!

Khatereh Nina Polous

Munich, 15.09.2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background and motivation

Geographic Information System (GIS) initially modelled geographical features and their relations independent of time. The main reason for this assumption was that most of geographical features maintain their identity as well as their location for relatively very long periods of time. Furthermore, early spatial information-collection approaches (mainly photogrammetry) concentrated only on capturing and recording such fundamental properties as identity and location. In addition, very high expenses usually made it impossible to repeat the capturing with a frequency that could possibly manage appealing change analysis and evaluation. Due to the perseverance of these essential and basic properties, representation of changes in time was not a preliminary consideration of Geographical Information Systems. Towards end of 80s and beginning of 90s, GIS society started to consider time in GIS and address the dynamic aspect of geographical features (Armstrong, 1988; Langran, 1992). This enhancement, for the first time provided the possibility of recording the history of objects with their attributes which could be used to predict possible future changes. The main emphasis nevertheless continued to be the geographical feature, and the temporal dimension was added as time stamps to keep track and evaluate different states of features. In this object change view – coined by Worboys, (2005) – objects have a unique identifier which sustains, but changes may happen to both spatial and non-spatial features. The conventional object-oriented view is clearly reflected in the ontologies that have actually dominated western thoughts at the very least since the Aristotle time (Rescher, 2008). This view sees the world as a collection of classified and identified objects, with specific properties and relationships. Nevertheless, there is an even increasing amount of work revealing that in numerous cases, the dynamic facets of geographical phenomena are vital to create useful explanatory, informative and predictive models. Hägerstrand, (1970) highlights the importance of time in human activities to assess the dynamic behavior of people in space, especially the motion of individuals in space and time. Miller, (2003) and Yuan, (2001) have exhibited this fact in their works on "transportation and urban analysis", and on analysis of physical phenomena, such as storms. Different researchers (such as Mark, (1998) and Miller, (2003)) have promoted the work of Hägerstrand, (1970) under the principal of geo-spatial lifelines. However, Hornsby and Egenhofer, (2000) deal with the object change view through the concept of identity-based change.

In general there are some main downsides when modeling changes with the object change view (Langran, 1992; Worboys, 2005):

- Relatively expensive computations and calculations are needed to detect and identify changes between snapshots, which is due to the fact that both snapshots need to be compared extensively.
- Developing or imposing rules for internal reasoning is challenging, since there is no understanding of the restrictions upon the temporal structure.
- No matter what is the size of changes, a full snapshot is produced at each time sequence. This replicates all unmodified information, which leads to storing huge amount of redundant information.

- And particular challenge in a case that the model only concentrates on objects' changes rather than a snapshot is related to establishing and also preserving the identity of an "individual object over time". Questions naturally arise as to "when and what change becomes so substantial that an object is no longer the same object".

Due to such restrictions in these models that merely include time stamps for managing versions and changes to the state of geographical attributes or locations, many researchers have suggested event-based models as an alternative solution (Claramunt and Thériault, 1995; Peuquet and Duan, 1995; Worboys and Hornsby, 2004; Worboys, 2005). Although early calls to maintain and preserve records of events and processes in order to understand dynamic behaviors go back to late 80s (Chrisman, 1998), "the realization of the event view owes much to new technologies that are now able to deliver a wide range and volume of spatiotemporal information" (Beard, 2006). Big repositories of information with high temporal resolution are created by environmental monitoring and sensor data streams to analysis occurring changes. Many temporal data streams with fine resolutions help in understanding how processes are working, which will be a basis for exploring cause and effect relations. In addition, considering that data is increasingly available, renewed impetus to develop tools and models for managing events and processes are needed (Beard, 2006).

In event-based models, change is the main concept that is modelled and change units are the primary items for analysis and evaluation. Claramunt and Thériault, (1995) define events as things which occur. Particularly they explain that processes cause changes in the state of objects, these changes reveal the outcome of the process and create events. Peuquet, (1994) defines an event as indicator of changes in a place or an object. Peuquet and Duan, (1995) refer to an event as a way to represent spatiotemporal manifestation of processes. Worboys, (2005) and Worboys and Hornsby, (2004) define an event as a happening that should be differentiated from a thing or continuant. They comment that the main weakness of snapshot models is the lack of an explicit representation for events and changes. Furthermore, they suggest that events are necessary to record the mechanism of change. In this approach the spatial dimension is dominated by the time dimension, because the sequence of events in time is essential (Beard, 2006). As the primary units of analysis and the dimensions are essentially different, new methods are required for modeling changes in objects, in ways that objects participate in events and relationships between events are modeled. In another word, as in object view, typically the globe is considered as some entirety of things; in event view the world is a totality of events (Galton, 2005). Galton, (2004) makes "the distinction between histories that are functions from a temporal domain to attribute values, or properties of objects, and chronicles that treat dynamic phenomena as collections of happenings".

The literature has very well expressed the motivations for an event-based approach. Events play a prominent role in various research areas such as physics, philosophy, psychology, linguistics, literature, probability theory, artificial intelligence, and history. William Sewell, a historian and political scientist perceives that societal structure is "a product of the events through which it has passed", furthermore he adds that "to understand and explain an event… is to specify what structural change it brings about, and to determine how [that change] was effectuated" (Sewell, 2005). This view could be discovered computationally through event-centered information models. The majority of Historical Geographic information System (HGIS) projects represent "human activity in terms of state, i.e. conditions of some geo-referenced entities in one or several temporal snapshots" (Grossner, 2010). "Geography and history are different ways of looking at the world, but they are so closely related that neither one can afford to ignore or even neglect the other" (Baker, 2003).

Events are perhaps the most extensive information container for dynamic geo-historical phenomena. In order to explain any event well enough, we should take into account its objective and results, its individual participants, its place in space and time, and its relationships to various other events. Indeed representing enough large number of events along all these dimensions may enable us to analyze and discover underlying social historical processes of the globe (Grossner, 2010).

Event-based modelling supports "representation of dynamic behaviors of geographical phenomena, hypothesis generation, scientific investigation of complex relationships, an ability to investigate causal linkages and associated entities with influences and underlying procedures" (Beard, 2006). Another practical usage of event-based models is utilizing available temporal information from monitoring and sensor networks. These data have remarkable value yet not completely utilized due to limitations in integrating and assimilating heterogeneous spatial and temporal information. However, most of the researches in this area are focused mainly on modeling events and relationships between them. Time information is included as an attribute for spatial entities or as an integral part of spatiotemporal objects. Furthermore, semantic information has actually not been much taken into consideration.

In this research, the concept of "mapping" moves beyond the principle of mapping objects. The main goal of this thesis is to present a conceptual model for managing geo-knowledge. This model can handle real world dynamisms such as the relationships amongst objects with each other, objects with events, events with events and the involving processes. This study uses a generic event-oriented perspective to implicitly represent causal relationships among different components of a Spatio-Temporal Geographic Information System. From this new perspective, the objects in space and time are considered merely as information elements of the events, which are connected to other event elements through internal or external processes.

There are different sources to collect required spatio-temporal information for further investigation such as photogrammetric data, satellite image, data collected through any kind of sensor and even field work. In this study, Volunteered Geographic Information (VGI) is used as it provides a unique opportunity to harness public's knowledge and their understanding about their surroundings directly; indeed we tap into the power of each individual. In addition, this information (specifically talking about spatial user generated content) is free and in their nature is georeferenced and timestamped. Furthermore, there are many types of spatio-temporal information which are difficult or even impossible to be collected through aforementioned methods, yet intelligent human beings can easily collect them. Making use of the unique opportunity offered by the concept of Volunteered Geographic Information (VGI), this study utilize VGI to provide the possibility of collecting and storing event related information elements in OpenStreetMap (OSM) platform.

Indeed, the introduction of the World Wide Web and particularly web-mapping technologies, which has been largely symbolized by the omnipresent Google Maps, has actually made digital mapping easily available to amateur individuals and users (Turner, 2006). The web 2.0 technology not only allows individuals to simply navigate maps over internet, users can generate and share their own geographic information, known as user-generated data or Volunteered Geographic Information (VGI) (Goodchild, 2007a). The term VGI has been commonly used to describe information with a spatial element, which are produced by volunteers. Goodchild, (2007a) describes VGI as the production of geographic information by mainly untrained volunteers, which is not restricted to conventional geographic identifiers such as trees and roads but to any kind of data that has a geospatial aspect. Tulloch, (2008) defines “VGI applications as those in which people, either individually or collectively, voluntarily collect, organize and/or disseminate geographic information and data in such a manner that the information used by many others”. Indeed VGI refers not only to a type of data but to the whole phenomenon of collaborative generation of spatial information. Other terms such as spatial user-generated content or collaboratively contributed geographic information have been used for the same purpose too (Bishr and Mantelas, 2008). Although, VGI is the most widely used by academia, it may not be the most accurate term. Because, information collection may not be volunteered at all, instead it may be facilitated or even gathered unknown to the volunteer (Sieber, 2007; Tulloch, 2008).

Elwood, (2008a) comments that VGI covers a vast range of topics from the generation of geographic features and attributes to arts and human rights. Hardey, (2007) remarks user generated geo-spatial data as one of the dynamic forces behind the revolution in experiencing the world. Furthermore, Goodchild, (2007b) and Elwood, (2008c) stated that VGI can be considered as a global patchwork of valuable and useful information, with space and time as its contextual glue. Finally, Goodchild, (2008b) Remarks that “the rapid growth of VGI in the past few years is one more step in a lengthy process that began almost two decades ago, and will likely continue for some time to come. It is one part of a fundamental transition as society redefines its vision of the role of public information in the early years of the 21st century”.

VGI has actually now obtained a genuine value and relevance in generation and dissemination of geographic information, and is not anymore considered merely as a hobby for amateur cartographers. VGI as an alternate and a complement to traditional sources of geographic information, “has greatly enriched our potential for characterizing the specificities of localities that central agencies either lack the resources, mandate or interest to collect and publish” (Hall et al., 2010). Moreover, the interconnection of numerous sources of information via web 2.0 technologies comprises a patchwork of knowledge that distributes very rapidly and broadly (Goodchild, 2007a). Consequently, VGI has actually become a distinctive source of information that can offer both researchers and people new insights on the world and the society we reside in.

One of the goals of VGI and GI is utilizing people, information, software and hardware in modeling the environment that we live in, in order to make better decisions. In 1998, Albert A. Gore the Vice-President of united state presented the idea of a Digital Planet to inspire and encourage researchers and scientists into creating a full-scale digital model of the planet earth that has connections to all information known to human (Wilson and Fotheringham, 2007). The advent of VGI based on the power of billions of potential human sensors with unique perceptions and worldviews, is indeed an important step to fulfill this goal. VGI can facilitate the connections between individual, educate them about their environment, influence the way they make decisions, and use their local knowledge for the good of all people and the planet. However to understand the nature of the VGI phenomenon and its impacts, it is essential to appropriately define VGI and its very different types. Furthermore, the volunteers, their motivations, and their abilities which consequently affect the quality, accuracy and credibility of the gathered data should be investigated.

By utilizing VGI, researchers are engaging volunteers for initiatives such as wildlife preservation, mapping invasive species, monitoring radiation degrees (Silvertown, 2009). The Christmas Bird Count project initiated by the National Audubon Society (established in 1900) to monitor changes in the population of birds is recognized as one of the pioneering examples in VGI. Aside from many examples related to biodiversity monitoring, VGI has been also used in remote sensing. For instance, Clark and Aide, (2011) used the Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT) to enhance the accuracy of reference samples via cross-verification of human as well as professional interpretations and analysis. This combination helped to successfully address problems associating with land change. As another example which illustrates the benefits of engaging volunteers with remote sensing imagery, we can refer to Geo-Wiki Project[1] that uses volunteers to improve the classification quality of remotely sensed global land cover maps.

Emergency management, disaster response and land management are other ways that communities and governments are using VGI (Haklay et al., 2014). As an example, during the Haiti earthquake in 2012, with the help of Ushahidi[2], volunteers contributed vital information such as trapped and injured people, damaged facilities and infrastructure, and closed or open roads. Moreover, VGI is used in enhancing the temporal accuracy of outdated information and datasets (Sui et al., 2013). As an excellent example of this we can refer to OpenStreetMap (OSM). The Canada-OSM Synergy Project which has been launched by Natural Resources Canada (NRCan), GIS data such as road networks are uploaded to OSM in order to be modified and kept up-to-date by citizens (Haklay et al., 2014). There are other mobile applications which have been created by municipalities such as the City of Calgary to enable residents in reporting specific (geo) concerns on bike paths, streetlights, snow and ice, water and sewer services, traffic signs, street and park lights, and road maintenance. City authorities utilize these information to reduce costs and expenses, and deploy resources where required.

Aside from previously discussed initiatives, the involvement of volunteers can be highly beneficial for diverse other fields such as in scientific research. Silvertown, (2009) notes “research funders such as the National Science Foundation in United States and the Natural Environment Research Council in United Kingdom now impose upon every grant holder to undertake project-related science outreach”. He suggests that this inclusion of volunteers in research can be an exceptional opportunity to spread out knowledge and educate the populace at the very same time. Bonney et al., (2009) showed how their study raised scientific understanding amongst their participants on ornithology, particularly because it related to bird watching. In addition, evaluating more than 230 projects by the UK Environmental Observation Framework (UK-EOF), it was concluded that volunteers provide information with a “high value to research, policy, and practice” (Tweddle et al., 2012).

[1] Geo-Wiki Project: Improving the quality of global land cover maps with volunteers – geowiki.org
[2] www.ushahidi.com

The underlying philosophy behind outreach is tapping into pools of individuals outside the organization that have not been traditionally included. Developing sophisticated open-source software is one of the best examples of the power of sheer numbers. At InnoCentive, many problems are solved by those who work on an entirely different domain from the field of the problem (Lakhani and Panetta, 2007).

VGI can indeed be considered as an organic development within cartography, an area that views the cartographers more of a group leader than a foot soldier (Crone, 1968). Under this scenario, the professional cartographer becomes an editor and a quality controller of gathered volunteered information that constitutes the main body of the spatial information. Filtering and controlling of VGI data has shown its usefulness, but (Flanagin and Metzger, 2007) comment that in general for VGI the "professional and scientific gate-keeping that usually filters and reviews data may not be present in sufficient forms and subsequently can lead to information which is prone to being poorly organized, out-of-date, incomplete, or inaccurate". Hence, some degree of caution must be given to the level to which VGI can be considered as the solution for more professional maps. Because, "most internet mapping users may lack sufficient cartographic training to manage or interpret the dynamic representation of geospatial information" (Tsou, 2003). Although this is a generalization of a wide variety of VGI contributors, it highlights one of the potential limitations of VGI. To improve the nature of the data, applications and tools should be designed in a way that can prevent using different terminologies and at the same time structure the gathered data (like tags).

Furthermore, current online mapping developments have actually not yet managed to handle most kinds of VGI which contain qualitative and vague information. Indeed, when it comes to deal with VGI specificities, little work has been done (Deparday, 2010). The phenomenal technological changes can substantially assist the idea of VGI to generate sophisticated local data comparable with available commercial and official datasets. However there are multiple challenges that make the integration of citizen-generated geospatial data with available official data a challenging and complex task. These challenges are for instance difficulty in assessing quality and credibility as well as redundancy of the gathered data as they have been generated by amateur users. The redundancy of the data comes from the fact that multiple contributors may gather similar information in parallel; especially when users use different terminologies and have different perceptions of a phenomenon in space. This heterogeneous user-generated information can substantially restrict usability of the data for decision-makers and especially citizens (Carver, 2003; Sieber, 2006). Because, expert users and researchers still have access to sophisticated analysis and visualization tools that can help them to discover meaningful patterns by exploring complex and multifaceted VGI content. This is not but the case for most of the amateur users. Hence, to assist novice and average users with special technical constraints who operate on web, visual-analytic tools should be adopted and redesigned to match the need of the average users, for instance by providing web-based tools to visualize the results instead of desktop-based application.

1.2 Hypothesis and research questions

As mentioned before, in this research the concept of "mapping" moves beyond the principle of mapping objects. The main goal of this thesis is to present a conceptual model for managing geo-knowledge. This model can handle real world dynamisms such as the relationships amongst objects with each other, objects with events, events with events and the involving processes. This study uses a generic event-oriented perspective to implicitly represent causal relationships among different components of a Spatio-Temporal Geographic Information System. From this new perspective, the objects in space and time are considered merely as information elements of the events, which are connected to other event elements through internal or external processes. In order to achieve the main goal of the thesis two main hypotheses were developed:

Hypothesis 1: If we develop an event-centric model for mapping, which is able to represent different components of events and their relationships, then Geo-Information Systems will be empowered to record and represent historical states of the globe along with its dynamic aspects to reach a geo-knowledge system.

Hypothesis 2: If we establish appropriate technological infrastructure for volunteers in VGI concept, then VGI can help to harness public's knowledge to build a geo-knowledge system.

In order to evaluate these two hypotheses some research questions are initialized. This study aims to answer the following research questions:

1. What is the concept of event mapping and which added value it brings into Geo information Science (GIS)?
2. How we can fulfill the concept of event mapping and move beyond the concept of object mapping?
3. How principle of Volunteered Geographic Information (VGI) can contribute in this move?
4. Which kind of technological infrastructure is needed to facilitate the contribution of volunteers for event mapping? How this development can help to structure unstructured data collection by contributors?
5. How to change role of volunteers from collection to analysis in order to increase spatial awareness of the crowd? And in which way it helps to move beyond Volunteered Geographic Information and reach Volunteered Geographic knowledge?

To sum up, the major contributions in this work are listed in the following:

- Move beyond the concept of object mapping
- Explore the concept of event mapping in order to build a geo knowledge system
- Propose an event-oriented modeling approach and utilize it to change the mapping process
- Utilize VGI to contribute in the concept of event mapping
- Store unstructured contributed data in a structured way in VGI
- Utilize VGI to develop a volunteered location based service
- Change the role of volunteers from merely data collectors to knowledge producers in order to increase the spatial awareness of the crowd

Figure 1-1 shows how this thesis is organized. Chapter 1 describes the background and motivation behind this research. Chapter 2 provides a review of the state of the art of the Volunteered Geographic Information and proposes a framework for VGI. Similarly chapter 3 presents literature opinion on the notion of event and event modeling, using this basis, a novel and generic event-oriented model for mapping events is developed in this chapter. Chapter 4 presents an event-centric framework for capturing and storing event information. This chapter focuses on the system architecture of the framework and four mechanisms that have been developed to collect event information. Chapter 5 emphasizes on first utilizing and enriching VGI through providing location based services using the collected information. Second, it explains a developed framework which aims at changing the role of volunteers from merely data collectors to knowledge producers. And finally, chapter 6 concludes the thesis and suggests further research work to extend the presented work.

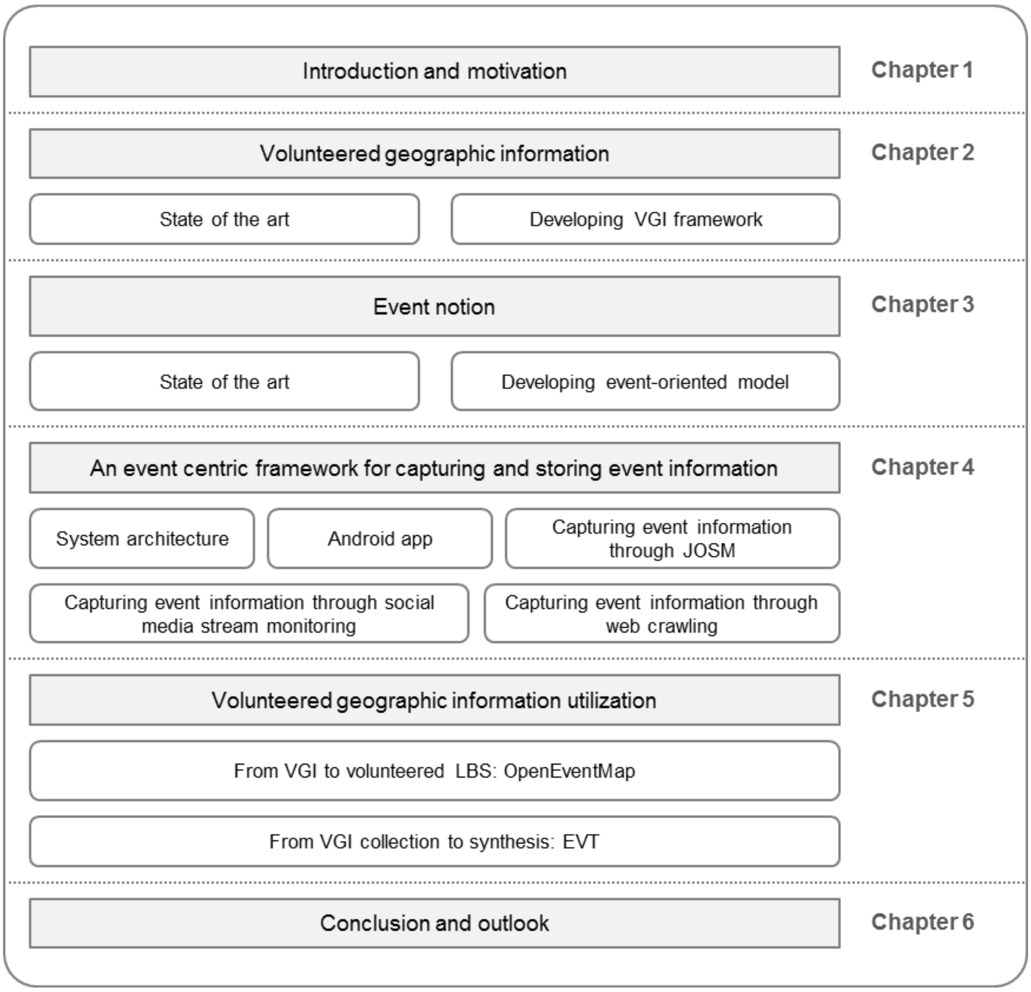

Figure 1-1: Thesis structure.

2 VOLUNTEERED GEOGRAPHIC INFORMATION

The very first wide public usage of the World Wide Web was the publication of static web sites that could be seen by anybody connected to the Internet net. In this realm it was only possible to get information on the Internet. Indeed a one-way information flow from facilitators such as municipalities or NGOs to the users (Rowe and Frewer, 2005). Although, enhanced web 1.0 applications enabled the facilitators to survey the public by allowing a flow of information from them, true interaction was never possible. The web technology evolved radically with the advancement of new design and programming techniques referred to as asynchronous JavaScript and XML (AJAX) that enabled the development of a more interactive and dynamic web (web 2.0) where a part of websites can be altered or refreshed without reloading the whole webpage, thus enabling individuals to interact dynamically with the display and consequently with each other (Mahemoff, 2006). Web 2.0 has actually transformed the way that individuals communicate, interact and share information with other individuals. Information flows on the web, and everybody can share different information pieces pertaining to specific areas on the planet. By putting together all discussed and shared pieces of information by Internet users, one can comprehend the characteristics and dynamics of the considered environments in the past as well as at present.

Tim O'Reilly coined the term web 2.0 in 1999 to describe the second generation of the World Wide Web that enable individuals to collaborate and share information online (Reilly, 2005). According to O'Reilly (Reilly, 2005), the core competencies of web 2.0 systems are:

- services, not packaged software, with cost-effective scalability,
- control over unique, hard-to-recreate data sources that get richer as more people use them,
- trusting users as co-developers,
- harnessing collective intelligence,
- leveraging the long tail through customer self-service,
- software above the level of a single device,
- lightweight user interfaces, development models, and business models.

Murugesan, (2007) describes web 2.0 as the second phase of internet's evolution: "harnessing the web in a more interactive and collaborative manner, emphasizing peers' social interaction and collective intelligence, and presenting new opportunities for leveraging the web and engaging its users more effectively". Web 2.0 focuses on data and contents, especially the capability to produce and interact with rich contents instead of being merely a consumer.

Web 2.0 envisions a web of individuals; in which everybody can create and share content and applications on the web. Applications are not any more single software programs on a computer that can be used and controlled only be one user; instead, applications are quickly accessible for everyone via Internet. This is possible only through utilizing light-weight and standard models and web development languages such as RESTful[3] and XML[4]. Consequently, an ever increasing number of public Application Programming Interfaces (APIs) has enabled individual developers and designers to call functions and data from numerous different sources. Considering that information is essential to web 2.0, web 2.0 systems flourish and grow on network effects: indeed, databases become richer as more and more people interact with them, applications end up being smarter as more individuals use them. As a result, such systems motivate users to involve more in sharing data. Web communities, hosted services, Social Network Services, forums, and blogs are examples of such web 2.0 based services.

In the context of geographic information, the expanding appeal of web 2.0 systems have created huge sets of openly available data, which consist of much explicit and implicit geographic information such as Volunteered Geographic Information (VGI), and social media data respectively. As the number of Internet users increases constantly, this huge amount of information increases as well. Furthermore, the disseminated information via social media shares ambient geographical information, because the feeds usually have geographical footprints such as the location where tweets are originated and geographic entities (Stefanidis et al., 2013). The enormous amount of created information in social media sites offers unique opportunities to understand and draw out valuable geographic information content. Both VGI and social network sites have actually been considered as crowd sources to gather geographic information from the internet, which have lately attracted a great deal of researches and investigates.

2.1 Crowdsourcing

The spreading of collaborative social media tools incorporated with the reach of the Internet has opened the possibility for people worldwide to share their knowledge, express their creativity and imagination, and make their voices heard in a way that has never been practiced (Parameswaran and Whinston, 2007). This has been further accelerated through affordable hardware tools and software programs that make it even possible for individuals to design, produce, and offer their own services and products easily. According to Tutty and Martin, (2009), individual users engage at different levels with the web 2.0 driven information. Initially, individual users involve with a social and collaborative environment where they receive and consume information in a unique and special way fitted to them. At the next level of engagement, individual users utilize the web 2.0 tools to collect and create knowledge based on their interest and passion, hence the meaningful relationships between the individuals as well as their technological environment increases. Ultimately, individual users will have the ability to manage the knowledge and information effectively, which consequently influences future social collaborations. Indeed, web 2.0 tools are not merely a set of tools that users may engage with, but it actively encourages a cyclic process of information acquisition, generation, management and dissemination (Tapscott and Williams, 2008).

[3] Representational State Transfer (REST) is a software architecture for designing distributed systems such as the World Wide Web.

[4] Extensible Markup Language (XML) is a markup language developed to structure, store, and transport information through a set of rules to encode documents in a format that is readable for both humans and machine.

An ever expanding number of public, private and commercial organizations are looking to "crowd" as a new possible source of innovation, knowledge, creativity, imagination, and productivity. The term generally described as "crowdsourcing," is a new technique to leverage online technologies to connect organizations and companies with individual contributors and new workforces around the globe. This term has been first coined in 2006:

"Crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers." In other words "the application of Open Source principles to fields outside of software" (Howe, 2006).

Doan et al., (2011) define this term as a system which "enlists a crowd of humans to help solve a problem defined by the system owners across a range of different industries". Different works have identified key dimensions of the multifaceted crowdsourcing concept, and explored the implications of models and theories from other domains in the area of crowdsourcing (Erickson, 2013; Lakhani and Panetta, 2007; Malone et al., 2009).

2.1.1 The power and value of crowdsourcing

We are just starting to comprehend the theoretical and practical implications of this emerging phenomenon. The increasing numbers of studies found in the literature shows that numerous organizations are trying to experiment with crowdsourcing concept (Allen et al., 2008; Brabham, 2010, 2008a; Chilton, 2009; Howe, 2008; Jouret, 2009). Both established organizations and new businesses have used it. Crowdsourcing has broadened significantly to a large range of topics covering commercial, industrial and scientific projects (Andriole, 2010; Bonabeau, 2009; Howe, 2008; McAfee, 2009). These include – among others – tagging images or documents (Barrington et al., 2009; Faymonville et al., 2009; Ho et al., 2009), developing creative designs (Brabham, 2010; Chanal and Caron-Fasan, 2008; Felstinerf, 2011), gathering distributed data (Sullivan et al., 2009), sharing knowledge and expertise (Bonabeau, 2009; Howe, 2008; Jana, 2009), solving complex problems (Albors et al., 2008; Archak, 2010; Bonabeau, 2009; Brabham, 2008b), and designing and developing evolutionary and revolutionary products (Bichler et al., 2010; Bretschneider et al., 2011; Di Gangi and Wasko, 2009; Jeppesen and Frederiksen, 2006; Poetz and Schreier, 2012). While each organization faces different needs and challenges, they leverage crowdsourcing as a way of getting work done. Indeed, the idea of crowdsourcing has expanded from merely collecting individually created work together to outsourcing contributions from a crowd, in order to complete a job executed previously by an employee or a professional contractor. The crowdsourcing concept has also shown its effectiveness in numerous scientific research subjects such as language translation (Ambati et al., 2010), usability studies (Heer and Bostock, 2010), and GIS (Goodchild and Glennon, 2010; Hudson-Smith et al., 2009).

Research studies show several organizational benefits and advantages can be drawn from crowdsourced initiatives such as creating the potential for organizations to reduce costs, increase their innovative capacity, and reduce products time to market (Enkel et al., 2009; Poetz and Schreier, 2012). Intangible benefits that companies report, include leveraging the crowd such as better understanding of possible new business opportunities, better realizing the way that outside perceives the organization (Jouret, 2009), and creating a culture based on innovation (Dodgson et al., 2006). Indeed, the ability of organizations to remain competitive has a direct relationship with their ability to access and leverage distributed sources of innovation (Poetz and Prügl, 2010), because the collaboration and partnership with people outside the borders of the firms can improve firms' innovation capabilities and capacities (Poetz and Prügl, 2010; Poetz and Schreier, 2012). By increasing connections to outside resources, additional sources of knowledge and creativity will be available for the firms, thus organization's innovative potential increases (Chesbrough, 2006, 2003; Jeppesen and Lakhani, 2010; Terwiesch and Xu, 2008). Researches have revealed that external sources have been able to solve 29.5 % of the previously unresolved problems by inside research and development laboratories (Lakhani et al., 2007).

Although, new emerging examples constantly highlight the wide range of usages and contexts for the crowdsourcing concept, at a very high level the theme for many of them is gathering information and making decisions. Howe, (2008) describes this increase in the use of information from wikis, "crowdsourcing" concepts, social network media, collaborative software platforms and many other web-based tools as a paradigm shift in the decision making process of companies; the emerging era of "Decisions 2.0". Indeed, true leveraging of crowd knowledge is not resumed in merely gathering information but utilizing the crowd to make decisions in order to solve today's demanding problems.

The decision-making process in the field of operations research consists of two high level tasks; 1) the generation of potential solutions, in which the problem is framed and a set of working assumptions are established and 2) the evaluation of alternative solutions in the first step (Bonabeau, 2009). Numerous human biases can negatively influence these two tasks to make optimum decisions (Myers, 2002). For instance, seeking information that confirms our initial assumptions or even keeping those assumptions in spite of contradictory evidences are two biases that prevent people from making reasonable decisions. Those biases, however, can be mitigated and reduced using three collective-intelligence approaches, namely; outreach, additive aggregation and self-organization (Bonabeau, 2009). In Table 2-1 some of these biases and their mitigations are listed. As can be seen in the Table 2-1 the diversity of viewpoints that can be achieved using collective intelligence approaches can deter many biases such as self-serving and belief-perseverance biases in the first phase, and pattern obsession and negative framing biases in the second phase of decision-making process.

The underlying philosophy behind outreach is tapping into pools of individuals outside the organization that have not been traditionally included. Developing sophisticated open-source software is one of the best examples of the power of sheer numbers. At InnoCentive, many problems are solved by those who work on an entirely different domain from the field of the problem (Lakhani and Panetta, 2007). Additive aggregation is a kind of averaging collected information from myriad sources. Here a balance between expertise and diversity is needed to reach meaningful decisions. Finally self-organization is a mechanism through which interaction among group individuals is enabled, which results in a whole that is more than the sum of its individual parts (Bonabeau, 2001).

Decision Making task	Biases in the Process	Collective intelligence approach to mitigate biases	Examples
Generation of potential solutions	1. Self-serving bias (seeks to confirm assumptions) 2. Social interference (influenced by others) 3. Availability bias (satisfied with an easy solution) 4. Self-confidence bias (believes prematurely to have found the solution) 5. Anchoring (explores in the vicinity of an anchor) 6. Belief perseverance (keeps believing despite contrary evidence) 7. Stimulation ("only knows a solution when seeing it")	1. Outreach to obtain diversity of assumptions 2. Additive aggregation to obtain independent participants 3. Outreach to obtain diversity of "easy" solutions 4. Outreach to obtain diversity of solutions 5. Outreach to obtain diversity of anchors 6. Outreach to obtain diversity of beliefs 7. Outreach and self-organization to obtain diversity of stimuli	Google, Affinnova, InnoCentive, Threadless, Bell Canada's I.D.ah!, ManyEyes, Swivel, Marketocracy, Goldcorp, Delicious, Digg, Procter + Gam- ble's Connect and Develop, Salesforce. com's Idea Exchange, Dell's IdeaStorm, Cajun Navy, Netflix's contest, blogs, wikis, Delphi method, lead-user tool kits, open-source software, "sousveil- lance," recommendation engines, support forums
Evaluation of potential solutions	1. Linearity bias (seeks simple cause-effect relationship) 2. Local versus global (confuses local and global effects) 3. Statistical bias (avoids statistical analysis) 4. Pattern obsession (sees patterns when none are present) 5. Framing (influence by presentation of solution) 6. Hyperbolic discounting (dominated by short-term effect) 7. Endowment bias (has aversion to risk or loss)	1. Self-organization to obtain nonlinear interactions 2. Self-organization to obtain nonlinear interactions 3. Additive aggregation to utilize law of large numbers 4. Additive aggregation and outreach to obtain diversity of pattern detectors 5. Additive aggregation to obtain diversity of influences 6. Additive aggregation to obtain diversity of time scales 7. Additive aggregation to obtain diversity of risk profiles	Digg, HSX, Zagat, "American Idol," Affinnova, Threadless, Intrade, Google, StumbleUpon, Bell Can- ada's I.D.ah!, Delicious, Mechanical Turk, Marketocracy, Salesforce.com's Idea Exchange, open-source software, Delphi method, information (or prediction) markets

Table 2-1: Using collective intelligence to make less biased decisions (Bonabeau, 2009).

There are many examples available that show additional value is created through interaction such as Wikipedia and Dig (Bonabeau, 2009). However a poor designed interaction mechanisms can have a reverse effect that destroys value rather than adding to it. Hence, there are several key issues that managers must consider before tapping into the crowd intelligence. Issues such as losing control over the crowd sourcing process and not keeping the right balance between diversity and expertise (Page, 2008) can drastically affect the desired outcome (Bonabeau, 2009). However, there are other vital factors that should be properly understood and analyzed before designing a successful crowd mechanism. One of the most important factors is the human factor of crowdsourcing initiatives; what motivates the people to engage, what is for them value, why they react different to incentives. This will help in its turn to design a community that is self-organized and self-controlled.

2.1.2 Participant's motivations and community

Reported motivation by participants who are engaged in crowdsourcing activities are very diverse such as financial drives, improving creative skills, finding potential jobs, community love, recognition by peers and the community, challenge, enjoyment, and learning (Antikainen and Vaataja, 2010; Archak, 2010; Brabham, 2008a, 2008b; Lakhani et al., 2007). Indeed, indicated motivations are a mix of both intrinsic and extrinsic incentives, which is similar to the motivations of the open source software program contributors (Hars and Ou, 2002; Lakhani and Wolf, 2003; Lakhani and Panetta, 2007). Extrinsic rewards are either direct such as financial rewards or indirect such as recognition within the community. Intrinsic rewards on the other hands are gained through engaging in the task such as challenges and enjoyment of working on the task, and passion for the subject. Different studies have also evaluated various types of intrinsic and extrinsic incentives, and examined which incentives are more powerful in driving participation such as for ideation and who are the most productive participants (Boudreau et al., 2011; Ebner et al., 2009; Fredberg and Piller, 2011; Johnson and Filippini, 2009; Leimeister et al., 2009; Muhdi and Boutellier, 2011).

Some studies indicate financial incentives as prime driver (Brabham, 2008b), whereas others show that intrinsic rewards have been more motivating than monetary ones (Lakhani et al., 2007; Leimeister et al., 2009). Hobbyists and lead users are most active contributors to crowdsourcing initiatives, however hobbyists have been in some cases more inspired by personal passion than by monetary incentives (Brabham, 2008a; Jeppesen and Frederiksen, 2006). While status and recognition by others and sponsoring-firm may enhance evolvement (Huberman et al., 2009; Trompette et al., 2008), it may likewise hinder less experienced individuals from participation (Archak, 2010). Recognition by the community and sponsoring-firm are both motivating but acknowledgement by sponsoring-firms may be sometimes even more encouraging than by others in the community (Antikainen and Vaataja, 2010). Different researches also analyze the relationship between social exchanges in participation (Füller, 2010; Wu and Fang, 2010), specify the optimum number of individual participants based upon incentive structures (Boudreau et al., 2011), and identify how best to incentivize the most productive individuals (Füller et al., 2011).

As studies show, individual contributors in crowdsourcing efforts are encouraged by a wide range of both intrinsic and extrinsic incentives. Furthermore, interactions with others have an influence on the level to which people participate in crowdsourcing activities (Dholakia et al., 2004; Halavais, 2009). This shows the importance of understanding the impact of different incentives on increasing participation in different contexts; some motivation factors may not be present in different conditions. Understanding the dynamics that influence participation is indeed crucial to implement crowdsourcing initiatives. Key is to find a balance between factors to motivate the crowd; however this requires understanding the make-up of the crowd, which ties directly to the role of community.

Community has been considered as an integral component of crowdsourcing across literature (Di Gangi and Wasko, 2009; Kozinets et al., 2008; Malone et al., 2009; Trompette et al., 2008; Verona et al., 2006), however it is questionable if any crowdsourcing group can be considered as a true community (Haythornthwaite, 2011). While some describe any group of working individuals on a specific task as a community regardless of their relationships and interactions (Feller et al., 2009; Trompette et al., 2008; Whitla, 2009), many consider interactions among individuals of a group inevitable for the existence of the community (Dholakia et al., 2004; Haythornthwaite, 2011; Lakhani and Panetta, 2007; McAfee, 2009; Wasko and Faraj, 2000). Haythornthwaite, (2011) supposes that there are lots of crowdsourcing groups that do not rise to the level of a true community. She believes that quality of the relationships and interactions between individual contributors defines a community, in which individuals share goals, standards, and are ready to openly reveal for the good of the entire community. However others believe that a community can be represented by a group of people with similar interests and skills (Trompette et al., 2008).

Community Type	**Contribution Type, Granularity and Authentication**	**Individual to Group Focus**	**Recognition, Reputation, Reward**
Crowd	• Atomistic, independent • Addressing uncertainty, explicit knowledge • Rule-based contribution • Delimited contribution attributes • Single form defined by authority/owner, authenticated by formula • Pooled interdependence	• Anonymous • History of contribution unnecessary • Open membership; low effort to enter • Two-tier hierarchy: authority, contributor • Independent, repetitive, discrete contributions	• Quantitative recognition mechanisms, e.g., contribution rate • Internally relevant to the individual application or the arena of contribution • Quantitative measures of contribution to product
Community	• Connected, revised, negotiated • Addressing equivocality, tacit knowledge • Negotiated contribution • Variable contribution attributes • Multiple forms defined and authenticated by group consensus, norms • Reciprocal interdependence	• Attributed • History of contribution important for group • Review, gatekeeping to join; high effort for membership • Multi-tier hierarchy: novice to expert, newbie to experienced • Continuing, contingent, norms-based contribution to product and process	• Qualitative recognition • Internally relevant, permeable to field of interest • Internal: judgments of contribution quality, expertise • External: judgment of contribution quality, expertise re field of interest • Peer review (qualitative) judgments of contribution to product and process

Table 2-2: Two spectrums of collaborative activity, adopted from (Haythornthwaite, 2009).

Haythornthwaite, (2011) distinguishes a community from crowd. Indeed there is a continuum between the two ends of the spectrum from crowds to communities (Haythornthwaite, 2011). “At one end are efforts that harness the knowledge and talents of many (relatively) anonymous individuals through online systems that aggregates discrete contributions into a whole. At the other end are communities, that meld, form, and define knowledge through the continued efforts of among a set of known participants” (Haythornthwaite, 2011). They have each different contribution, participation, aggregation and evaluation patterns. Table 2-2 lists three dimensions of collaborative activities that shows the differences between “lightweight” participation (crowd) and heavyweight participation (community) (Haythornthwaite, 2009).

At the crowd end of the spectrum, contribution is simple and uncomplicated, with simple rules, coordinated with a pooled of similar interdependence contributions; at the other end of the spectrum, contributions need higher learning and are assessed by other individual participants in a peer review procedure. The second dimension – individual to group focus – influences the extent to which the organizations can, or must rely on the attention that individual participants offer to other individual contributors and the contributions of others. These dimensions set the basis for the third one: the recognition, reputation, and reward (Haythornthwaite, 2011). This insight clearly shows the cruial role of community in crowdsourcing initiatives. Indeed understanding the context and conditions of crowdsourcing initiatives helps us to better understand the motivation of individual participants, and will help to design appropriate "self-organizing" mechanisms (Bonabeau, 2009) to control and evaluate crowdsourcing results.

2.2 Geo-information Science Evolution

2.2.1 Public Participation GIS

In 1990s, considering different aspects of Geographic Information Systems (GIS) such as social, ethical, political and societal brought about vigorous debates and discussions among researchers in the field (Craig et al., 2002; Pickles, 1995). Among the primary concerns was the elitist and exclusive facet of the GIS technology that limited the access to spatial information to few ones in the society who can afford expensive hardware tools and software programs as well as costly geo-data (Pickles, 2005). Another essential point of debate between critics of GIS (Lake, 1993; Taylor, 1990) and supporters of GIS (Openshaw, 1991) was that GIS is based on positivist assumptions by using information that represents simple facts rather than more complex and multifaceted knowledge. Critics believe that by representing just single perspectives and reducing societal processes into points, lines, areas, attributes, and features multiple nuanced geographical realities can be overlooked (Sieber, 2006). Furthermore various other types of information that can have essential geographical dimensions such as past history, emotions and feelings, sacredness, and meaning for the local people are difficult to integrate into GIS data models which tend to only focus on the numerical tractable aspects of issues and problems (Deparday, 2010). The problem becomes more acute when considering the limited number of "official" sources of data that represent only a general and dominant view point of the reality, so different views and interests of other communities and their local knowledge can be easily undermined and marginalized (Pickles, 2005). This local knowledge is mainly lost either in the simplification process to fit into data models or is not at all appealing financially or objectively for the organization that generates the datasets. There are various definitions for the local or indigenous knowledge such as "unique, traditional, local knowledge existing within and developed around the specific conditions of women and men indigenous to a particular geographic area" (Grenier, 1998), or "value-based and traditionally intangible information" (Sieber, 2006), or finally as "knowledge that is unique to a given culture or society" (Warren, 1991).

These shortcomings in the data quality along with asymmetrical access to GIS technology as well as information constitute the base of the critiques who considered GIS as an anti-democratic technology which strengthens already existing departments in society by sustaining a mere top-down and exclusive approach toward decision-making where most citizens cannot participate (Pickles, 2005). Hence in the 1990s, academics, researchers and practitioners engaged in a research procedure that entailed new techniques, methods, practices, as well as tools to resolve the critiques in a constructive way. GIS-2, critical GIS, Community-integrated GIS (CiGIS), Participatory GIS (PGIS), or Public Participation GIS (PPGIS) are different suggested terms for this new field of research, however PPGIS is the most widely accepted term (Weiner and Harris, 2008). Two core principals of PPGIS, which are critical to democratize GIS within communities, are: disseminating GIS technology and data to marginalized groups such as grassroots movements and community-based organizations, and integrating communities' local knowledge in the GIS datasets (Carver, 2003; Sieber, 2006). By ensuring an easier access to GIS tools and technology with limited resources, PPGIS aimed to empower these communities through enhanced possibilities ranging from the merely exploration of their environments to involvement in formal decision-making procedures. PPGIS has been used to address diverse issues in very different contexts (Craig et al., 2002) such as; managing conflicts and disputes among stakeholders and increasing collaboration (Balram, 2006; Kyem, 2004; Weiner and Harris, 2003), neighborhood revitalization (Elwood, 2002; Ghose, 2001), land usage and planning management (Bojórquez-Tapia et al., 2001), environmental and ecological management (Evans et al., 2004; Jankowski and Nyerges, 2001), indigenous and native territory claims (Dana, 2008).

Nonetheless it is necessary to keep in mind that the idea of participation does not describe just a solitary homogeneous approach. To recognize who profits from accessing GIS technology and data and also why, as well as exactly what are the proper methods, it is essential to recognize the participatory procedure. To do so, vital questions have to be explored such as a) exactly what is meant by participation? as well as b) who is the public? (Dunn, 2007; Schlossberg and Shuford, 2005). Schlossberg and Shuford, (2005) provide a thorough review of various ways to identify the participation process. A typical approach is to utilize differing levels of public participation (defined initially by (Arnstein, 1969) as a ladder, which has been adjusted by many authors such as (Carver, 2003) Figure 2-1. The level of participation and responsibility changes between the two extremes of having a passive role or participating in final decision making rises. However, the actual empowerment of the public necessitates also an equal access to relevant information (Carver, 2003). As defined by (Schlossberg and Shuford, 2005), the public, so-called PPGIS participants or stakeholders, typically consists of citizens who belong to one or some of the three following categories: "those affected by a decision or program" such as neighborhood residents that may be also represented by non-governmental or community-based organizations (NGOs or CBOs), "those who can bring important knowledge or information to a decision or program" such as developers, and scientific groups, and "those who have power to influence and/or affect implementation of a decision or program" such as municipal or regional government officials.

To enhance the success rate and the impacts of PPGIS initiative, it is important to appropriately establish the kind of involvement as well as the public that is needed for a specific goal, with different social and cultural context. Rowe and Frewer, (2005) present a comprehensive list of participation methods. But integrating local knowledge in GIS needs better understanding of participants' skill levels, interests, and cultural backgrounds. Various studies have discussed different steps in incorporating local perspectives into official and professionally generated GIS data, namely how to gather, use and integrate local knowledge in GIS data models (Dunn, 2007; Hall et al., 2009; Sieber, 2006). Although, integration of local insights and knowledge in GIS data may help to enhance the accuracy, reliability and completeness of official data (Carver, 2003), adjusting GIS to include such vague information or reorganizing the local information with multiple perspectives to conform to GIS models are challenging problems if at all possible (Sieber, 2006). Because GIS is originally designed to handle and represent simplified factual information, it is not yet suited for information which contains several perspectives and ambiguity.

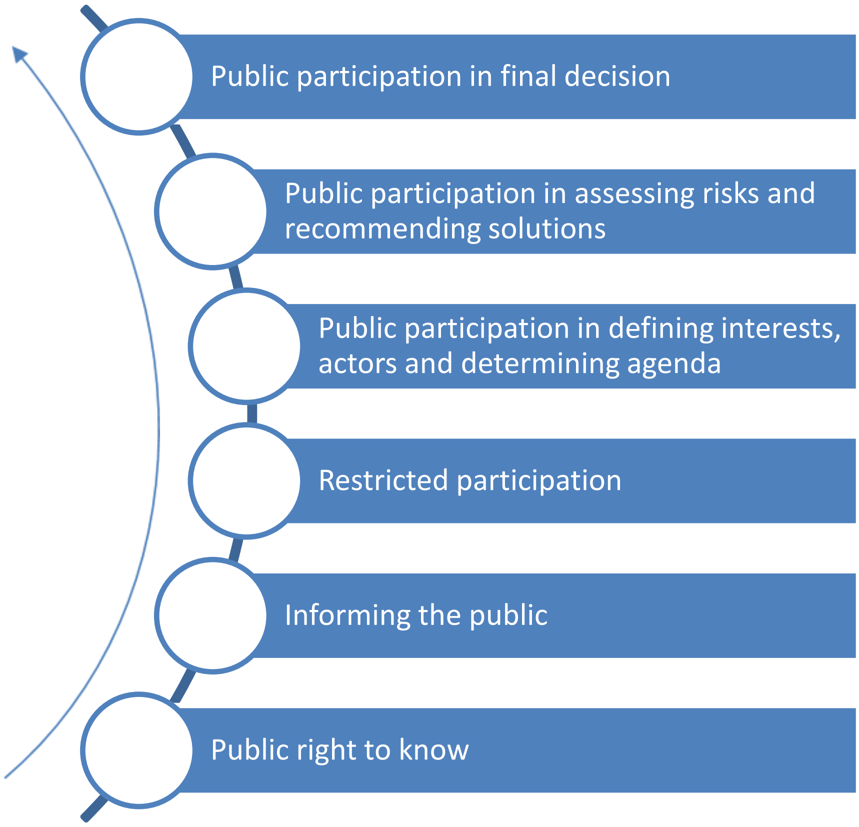

Figure 2-1: The public participation ladder (Carver, 2003).

Craig et al., (2002) present various case studies about collecting different types of local knowledge in several PPGIS efforts. Nevertheless, as a result of technological, institutional, and political issues, real empowerment via using local knowledge has been generally restricted (Weiner and Harris, 2008). Major technology-driven changes in GIS practice such as increasing number of web-based tools that are generally simpler to use and allows individuals and teams to be involved from anywhere anytime has revolutionized collection and effective use of local insights and knowledge (Carver, 2003; Sieber, 2006). Indeed the evolution of the web technology to web 2.0, has offered many opportunities to researchers to reach and involve a bigger audience of citizens via multiple crowdsourcing mechanisms such as Volunteered Geographic Information (VGI) that aid in producing valuable spatial and non-spatial User-Generated Content (UGC).

2.2.2 Web 2.0 and the dawn of Neogeography

Evolution of web technology to web 2.0 was in the same direction as efforts in the PPGIS field to make GIS more available and suitable for public involvement. Table 2-3 lists some differences between GeoWeb 1.0 and GeoWeb 2.0 technologies (Maguire, 2007). Indeed many of the GeoWeb 2.0 core ideas listed in Table 2-3 are crucial for the further evolution of the GIS and PPGIS areas. Web 1.0 technologies allowed simple participation types corresponding to the first two levels of the public participation ladder in Figure 2-1 on the web. Web 2.0 technology however acts as a platform, upon which different software can be developed as a service (Reilly, 2005). It has facilitated the emergence of diverse business models, practices and software services that are readily available on the web and fully functional via a simple web browser. This significantly lowers the acquiring costs for users who do not have to purchase costly software and hardware any more, consequently the access to technology broadens significantly. Furthermore, the “architecture of participation” in web 2.0 unlike web 1.0 that is designed for publishing, aims to facilitate the utilizing of collective intelligence (Tapscott and Williams, 2006) hence allowing a full interaction between the facilitators and public as well as between the public members themselves (Hall and Leahy, 2011). This is very important to reach a higher participation level of the public as shown in Figure 2-1, which underlies the collection of local knowledge. Finally, web 2.0 applications are developed from “a network of cooperating data services” (Reilly, 2005). This important aspect allows websites to directly use any number of remote information sources without duplicating them. For example, maps, calendars or weather forecasts from different web service providers can be seamlessly integrated in your interface.

GeoWeb 1.0	**GeoWeb 2.0**
Static	Dynamic
Publishing	Participation
Producer-centric	User-centric
Centralized	Distributed
Close-coupling	Loose-coupling

Table 2-3: Some differences between GeoWeb 1.0 and GeoWeb 2.0 (Maguire, 2007).

These technologies and concepts have significantly helped to democratize GIS tools and spatial information by minimizing unequal access to GIS data and technology. One of the main implications of web 2.0 technologies for GIS and PPGIS is that a widely distributed network of geospatial data users and providers can cooperate remotely over the internet. Web 2.0 has also extensively facilitated other facets of GIS such as open source GIS software development by many remote contributors or sharing computing power to solve time and resource demanding problems through a big grid of individual computers (Sui, 2008).

Web 2.0 technology has dramatically increased the use of GIS technology and data in the society. The popularization of GIS originally started with the omnipresent Google Maps and Yahoo Maps and their Application Programming Interfaces (APIs) that allow users to create tailored mapping tools (Gibson and Erle, 2006). Indeed complex cartography and GIS methods and techniques are not anymore inaccessible for the untrained public to create or use spatial information. Consequently, the Internet is loaded with maps created for diverse topics by amateur cartographers. This phenomenal wave is widely called Neogeography (Haklay et al., 2008; Turner, 2006).

Utilizing web 2.0 technologies can provide scientists an “ideal combination of scientifically sound, high-quality information that is imbued with experiential insights from a multitude of individuals” (Metzger and Flanagin, 2011). In this web 2.0 environment that maps are used as a participation platform, Eisnor, (2006) coined the term Neogeography: “A socially networked mapping platform which makes it easy to find, create, share, and publish maps and places”. Turner, (2006) clarifies this further:

"Neogeography means 'new geography' and consists of a set of techniques and tools that fall outside the realm of traditional GIS, Geographic Information Systems. Where historically a professional cartographer might use ArcGIS, talk of Mercator versus Mollweide projections, and resolve land area disputes, a neogeographer uses a mapping API like Google Maps, talks about GPX versus KML, and geotags his photos to make a map of his summer vacation. Essentially, Neogeography is about people using and creating their own maps, on their own terms and by combining elements of an existing toolset".

HousingMaps.com is one of the first Neogeographic maps, or "mashups" in the world. These new "mashups," led to a huge increase in the propagation of GIS(s). The usage of term mashup can be traced back to the first mapping mashups such as Housing Maps and ChicagoCrime (Turner, 2006). Miller, (2006) defines mashups as "new services built from the code and functions of two or more different, sometimes even disparate, projects". Floyd et al., (2007) remark that the utilized technologies in mashups are not new or necessarily innovative; however "what is innovative is how mashups are being widely used for the rapid realization of creative ideas which would be too time consuming or expensive". On the other hand, Wilson, (2009) believes that from a GIS perspective "mashups elude our traditional ways of knowing and seeing".

Although, there are many different positive and negative reflections on neogeographic mashups; many research studies have remarked different advantages and high quality of neogeography products (Elwood, 2008c; Goodchild, 2010). Furthermore, similar to other geographic systems, the quality and fitness for purpose of neogeographic systems should be assessed to address user concerns such as accuracy and completeness (Coote and Rackham, 2008; Goodchild, 2008a). Goodchild, (2008b) comments that accuracy issue is similarly common in volunteered as in professional information.

2.3 Volunteered Geographic Information

"Throughout most of the history of cartography, maps have been used by elite groups, to control and administer people and places" (Pickles, 2004). Nevertheless, readily available GIS(s) dispersed through the internet have marked an end to this era. Indeed, maps and mapping have undergone a substantial democratization in their accessibility, usage and appeal (Sui, 2008). The current geospatial revolution as well as advancement in web-based and online mapping technologies has facilitated the emergence of a relatively new domain referred to as Volunteered Geographic Information (VGI) (Goodchild, 2007a). VGI builds on previous participatory techniques, yet gets to a much larger audience, indeed to the broadest sensor network available today. Empowering 7-billion individuals globally can create a revolution in volunteer geospatial content that can be used for multiple purposes (Goodchild, 2007b). Additionally, VGI offers interesting collaboration possibilities between volunteers and geospatial researchers due to the advancement of interactive web mapping systems. Instead of relying merely upon a limited team to address complex and intricate problems, or collect geospatial information, researchers now have the ability to develop an open-ended and flexible system to engage volunteers from all over the world. However, VGI could possibly be considered as pure VGI only in its early days, because many projects such as OpenStreetMap (OSM) consisted of just information from amateur volunteers (Haklay and Weber, 2008). Nevertheless, within few years organizations such as the United States mapping agencies TIGER started to integrate their open data into VGI data sets (Black, 2007).

The term Volunteered Geographic Information (VGI) has been commonly used to describe information with a spatial element, which are produced by volunteers. Tulloch, (2008) defines “VGI applications as those in which people, either individually or collectively, voluntarily collect , organize and/or disseminate geographic information and data in such a manner that the information used by many others”. Indeed VGI refers not only to a type of data but to the whole phenomenon of collaborative generation of spatial information. Other terms such as spatial user-generated content or collaboratively contributed geographic information have been used for the same purpose too (Bishr and Mantelas, 2008). Although, VGI is the most widely used by academia, it may not be the most accurate term. Because, information collection may not be volunteered at all, instead it may be facilitated or even gathered unknown to the volunteer (Sieber, 2007; Tulloch, 2008).

Goodchild, (2007a) describes VGI as the production of geographic information by mainly untrained volunteers, which is not restricted to conventional geographic identifiers such as trees and roads but to any kind of data that has a geospatial aspect. Elwood, (2008a) comments that VGI covers a vast range of topics from the generation of geographic features and attributes to arts and human rights. Hardey, (2007) remarks user generated geo-spatial data as one of the dynamic forces behind the revolution in experiencing the world. Furthermore, Goodchild, (2007b) and Elwood, (2008c) stated that VGI can be considered as a global patchwork of valuable and useful information, with space and time as its contextual glue. Finally, (Goodchild, 2008b) Remarks that “the rapid growth of VGI in the past few years is one more step in a lengthy process that began almost two decades ago, and will likely continue for some time to come. It is one part of a fundamental transition as society redefines its vision of the role of public information in the early years of the 21st century”.

VGI systems have three core components; volunteers, sensors and Geoweb applications. Volunteers are amateur or non-professional neogeographers (Turner, 2006), who participate in research studies by gathering geographical information that is not discernable from an airborne, satellite, or any type of physical sensors based on their experience and local knowledge (Goodchild, 2007a). Sensors are participants and volunteers who use their own five senses (Goodchild, 2007a). These volunteers collect and share geographic information passively, actively, or through a hybridized approach (Feick and Roche, 2013). And finally the Geoweb application allows volunteers through a user interface (UI) to complete necessary tasks for gathering geographic content such as adding map features, geo-tagged media, and location-specific or attribute metadata (Goodchild, 2007a).

VGI has actually now obtained a genuine value and relevance in generation and dissemination of geographic information, and is not anymore considered merely as a hobby for amateur cartographers. VGI as an alternate and a complement to traditional sources of geographic information, “has greatly enriched our potential for characterizing the specificities of localities that central agencies either lack the resources, mandate or interest to collect and publish” (Hall et al., 2010). Moreover, the interconnection of numerous sources of information via web 2.0 technologies comprises a patchwork of knowledge that distributes very rapidly and broadly (Goodchild, 2007a).

Consequently, VGI has actually become a distinctive source of information that can offer both researchers and people new insights on the world and the society we reside in.

VGI can indeed be considered as an organic development within cartography, an area that views the cartographers more of a group leader than a foot soldier (Crone, 1968). Under this scenario, the professional cartographer becomes an editor and a quality controller of gathered volunteered information that constitutes the main body of the spatial information. Filtering and controlling of VGI data has shown its usefulness, but Flanagin and Metzger, (2007) comment that in general for VGI the "professional and scientific gate-keeping that usually filters and reviews data may not be present in sufficient forms and subsequently can lead to information which is prone to being poorly organized, out-of-date, incomplete, or inaccurate". Hence, some degree of caution must be given to the level to which VGI can be considered as the solution for more professional maps. Because, "most internet mapping users may lack sufficient cartographic training to manage or interpret the dynamic representation of geospatial information" (Tsou, 2003). Although this is a generalization of a wide variety of VGI contributors, it highlights one of the potential limitations of VGI.

One of the goals of VGI and GI is utilizing people, information, software and hardware in modeling the environment that we live in, in order to make better decisions. In 1998, Albert A. Gore the Vice-President of united state presented the idea of a Digital Planet to inspire and encourage researchers and scientists into creating a full-scale digital model of the planet earth that has connections to all information known to human (Wilson and Fotheringham, 2007). The advent of VGI based on the power of billions of potential human sensors with unique perceptions and worldviews, is indeed an important step to fulfill this goal. VGI can facilitate the connections between individual, educate them about their environment, influence the way they make decisions, and use their local knowledge for the good of all people and the planet. However to understand the nature of the VGI phenomenon and its impacts, it is essential to appropriately define VGI and its very different types. Furthermore, the volunteers, their motivations, and their abilities which consequently affect the quality, accuracy and credibility of the gathered data should be investigated.

2.3.1 VGI conceptual frameworks

Descriptive efforts to characterize and define VGI, its implications and effects are crucial to arrange future research studies on VGI (Elwood, 2008b). Indeed, VGI includes such a broad range of projects, and information that although specific strategies might be useful in some circumstances, yet they may not be appropriate in various other situations. Hence, a comprehensive framework is essential to investigate different challenges and their associated remedies based on the various kinds of VGI. There are several ways to describe and characterize different VGI phenomenon depending upon the aspects and facets that a study wants to concentrates. These include diverse characteristics such as the nature of the volunteered data, the purpose of the contribution, mechanisms that have been used to gather data and the approach to utilize these data.

The VGI literature describes and characterizes VGI through various elements. However efforts of explaining and identifying the broadest landscape of VGI through a single framework has not yet reported in the VGI literature. Frameworks related to VGI originate mainly from PPGIS and neogeography literature.

Turner, (2006) presents a framework for Neogeography, in which Neogeography brings different tools and technologies for gathering and showing location data together into a "GeoStack" Figure 2-2. "The GeoStack is a collection of tools and mechanisms that together cover all parts of collecting, gathering, and sharing location information. It enables a GPS system to capture a waypoint and eventually have other users around the world view and comment on that waypoint" (Turner, 2006). This framework has been developed based on the successive steps of VGI data management from capturing data by volunteers to its usage by an end-user. Although, the GeoStack framework provides a good summary of several different tools and technologies utilized in the production process of VGI data in the context of web 2.0, one of its downsides is the need for regular update of the framework. Because used tools and technologies in the framework are very rapidly advancing and evolving.

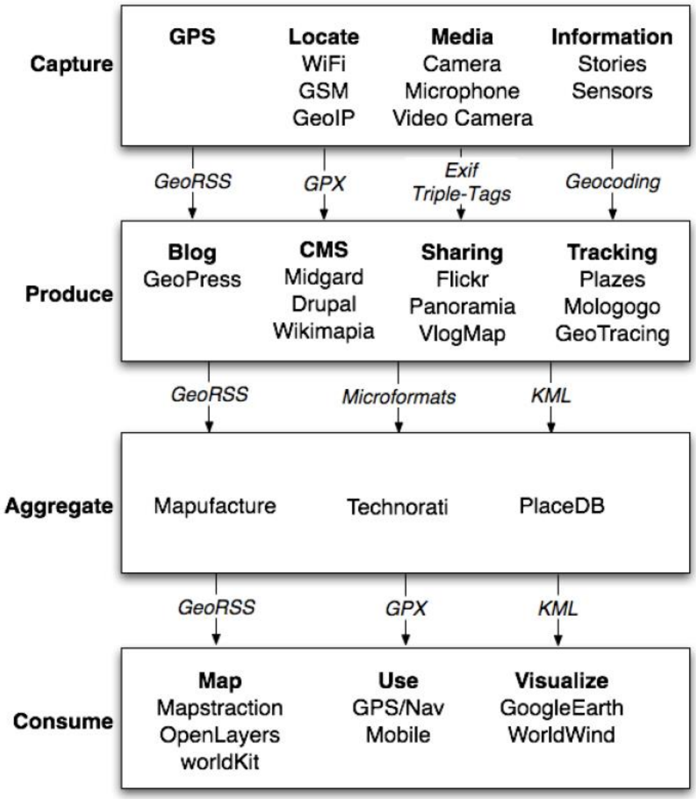

Figure 2-2: The GeoStack framework (Turner, 2006).

Schlossberg and Shuford, (2005) present numerous PPGIS frameworks that examine the participation of "the public", especially the type of the participations and their nature. Turkucu and Roche, (2008) presented a framework, which describes PPGIS based on six leading characteristics (Figure 2-3); public involvement, employed technology, data, and the interaction between technology and users. Considering the fact that Materials, Mechanisms and Public most of the time are mixed, only three axes remain (Turkucu, 2008); Software, Public Involvement and Data. These parameters do not allow characterization of many VGI elements in the recent VGI context, because todays VGI approaches utilize primarily web 2.0 technologies and Net with a high interaction between the individuals and the tools.

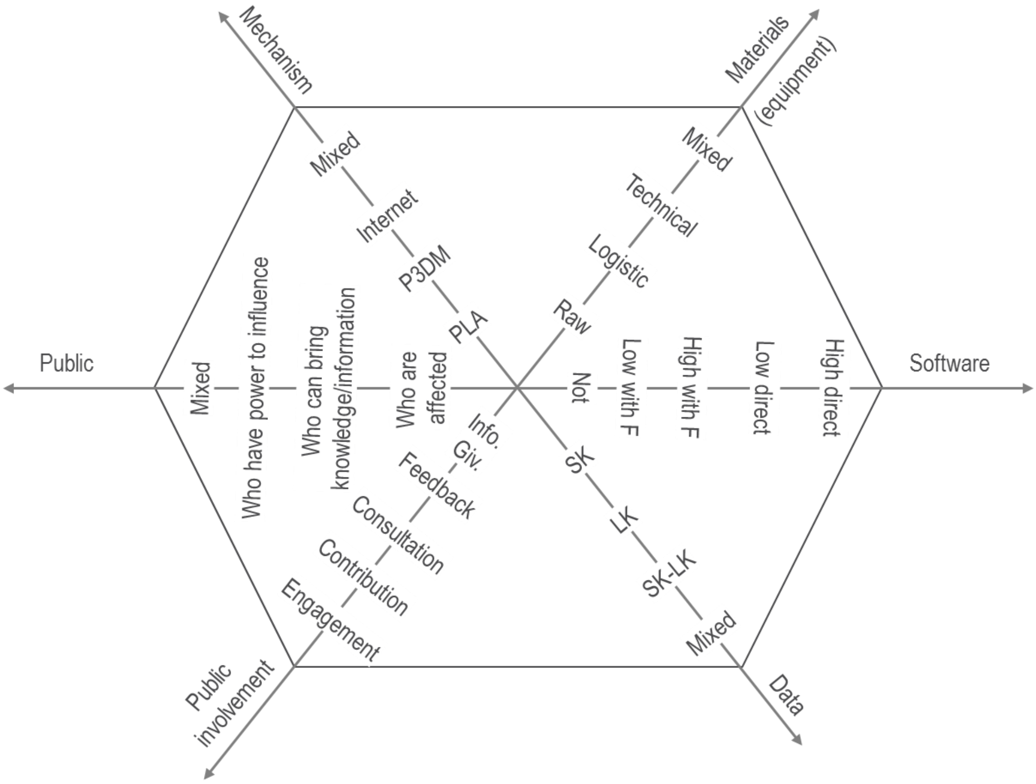

Figure 2-3: Leading characteristics of PPGIS (Turkucu and Roche, 2008).

The literature highlights various classification systems that have been used to distinguish between different VGI efforts. These classifications emphasize mainly those aspects of VGI that are critical for their presented study and do not capture the entire complexity of the VGI phenomenon. Coleman et al., (2009) present a set of models that focus mainly on characterizing the spectrum of contributors, contribution motivations, characteristics of use, and the institutional requirements. However these models cannot be used as an all-encompassing framework due to the absence of connection and integration between the produced models. For understanding quality issues of VGI, Cooper et al., (2011) presented two dimensions for VGI; "the continuum of responsibility for determining the specifications for the data" ranging from a user through to an official data custodian with tightly controlled specifications. The second dimension was "classifying data as base" ranging from base data to points of interest. Parker, (2012) suggests two rating scales for measuring both quality control measures and the degree of objectivity. Building on the rating scales and evolving from the Cooper et al., (2011) approach, he presents a framework to categorize neogeographic products, within which, neogeographic products are assessed based on the rating criteria.

The most comprehensive conceptual framework for VGI has been presented by (Budhathoki, 2010; Budhathoki et al., 2010) as shown in Figure 2-4. He proposes a three-tiered conceptual framework for VGI for classifying VGI-related elements. The framework has three arenas ''Motivation'', ''Action and Interaction'' and ''Outcome'', and each arena has different sub-arenas. The motivational arena refers to individual people and their inspirations to contribute. The action and interaction arena deals with the decision-making process and how people contribute. The outcome arena consists of contribution and evaluative criteria (Budhathoki et al., 2010).

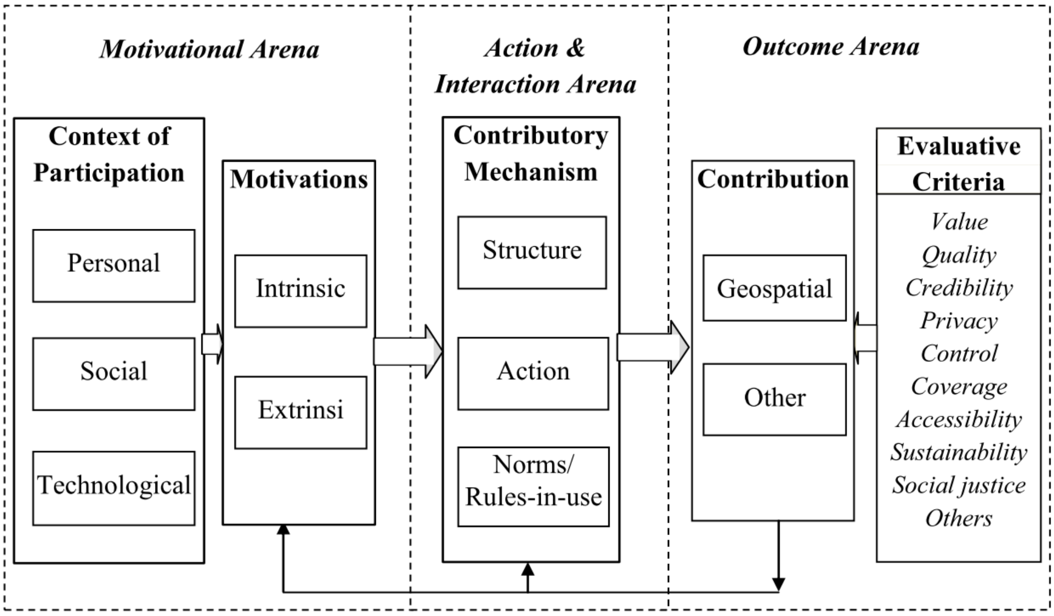

Figure 2-4: A conceptual framework for VGI (Budhathoki et al., 2010).

While the recommended arenas offer a good overall framework to analyze VGI, more comprehensive conceptual models for evaluating different sub-arenas are missing Rehrl et al., (2013). Furthermore, the outcome arena does not cover the utilization of the VGI information. In order to fill this gap, here a conceptual framework for VGI is proposed and is used to investigate different aspects of VGI. It can be used to systematically explore various aspects of VGI and analyze the relationships between different elements and processes of VGI.

2.3.2 Developing a conceptual framework for VGI

A framework is an analytical skeleton which contains a set of logical and rational building blocks as well as their interconnections (Hess and Ostrom, 2007). By giving a broad view, it assists us in analyzing problems more holistically and offers a valuable tool for systematic and methodical investigation of static and dynamic scenarios and circumstances. In order to distinguish and identify VGI, it is essential to specify first a set of parameters and criteria that allows evaluation of the similarities and differences across the VGI panel, and ultimately to determine the major sorts of VGI. These parameters are coming mainly from the available models and frameworks in the PPGIS, neogeography and VGI fields as well as from critical analysis of the VGI landscape. The parameters are organized in four categories; enablers, context, mechanisms, and utilization (Figure 2-5). Each category has in its turn various parameters that are used to characterize and identify different VGI types.

Enablers – unlike other categories – are not a set of parameters or criteria for describing different characteristics of VGI. VGI inherits indeed many of its features from scientific and technological enablers such as web 2.0, GIS and PPGIS that have been extensively described in the previous sub sections. To fully understand VGI, we should admit and understand the relationship of enablers to their phenomenal offspring VGI. Hence, the basis of the proposed framework has been shown as scientific and technological enablers. Context describes the essential distinctions in the nature of the various types of VGI. This category represents various VGI continuums. Mechanisms represent different approaches and tools that facilitate contribution. And finally utilization describes last phase of the VGI cycle; data utilization and synthesis that leads to making decisions based on the VGI information.

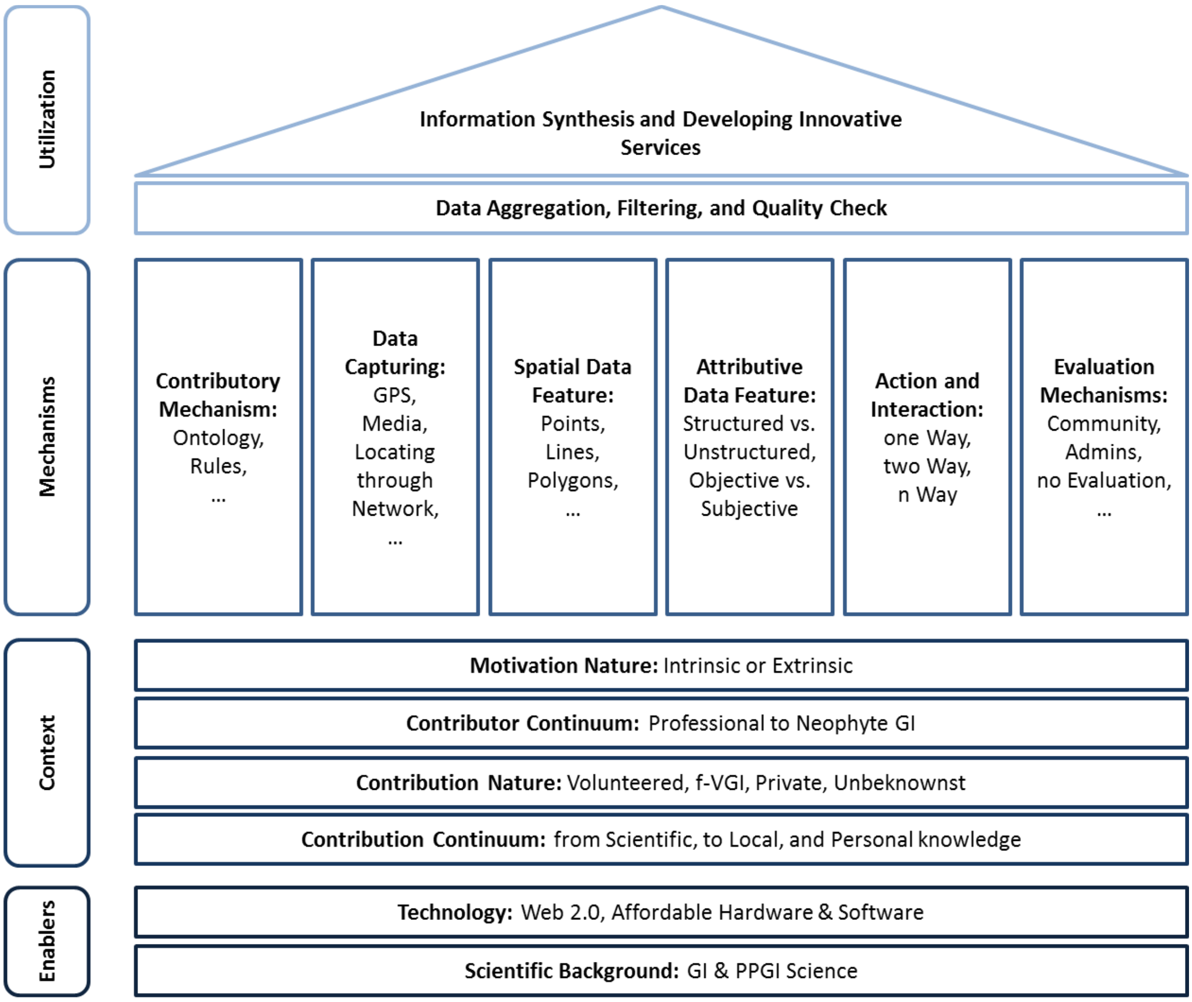

Figure 2-5: VGI conceptual framework.

2.3.3 Context of VGI

The context of VGI presents different fundamental natures of VGI that significantly influence the contribution process, and the nature and quality of the gathered data. Different continuums elaborated in this section, provide an important basis to structure diverse VGI types. The context of VGI consists of four main continuums and natures: a) contribution continuum, which varies from scientific knowledge to personal knowledge, b) contribution nature that addresses different facet of the contribution itself such as volunteered or unbeknownst, c) contributor continuum that ranges from professional to amateur contributors, and d) the motivation nature of volunteers that explains intrinsic and extrinsic motivations.

2.3.3.1 Contribution continuum

Scientific knowledge (SK) within the context of VGI describes the contributed knowledge by volunteers which is either scientific by nature or describes the world in a quantitative, measurable and scientific way. According to this definition, mapped objects, street networks, resident's densities, and soil, plant and other classifications are considered as scientific. Local knowledge (LK) ranges from conventional GIS to qualitative understandings and opinions that belong to a specific perspective.

These two types of knowledge are adopted from leading characteristics of PPGIS (Turkucu and Roche, 2008). Finally, the personal knowledge (PK) is at the other end of knowledge spectrum. This knowledge is generated to be shared with those in personal circles such as friends, relatives and co-workers and is for personal purposes such as georeferenced pictures and recommendations about point of interests or events. Deparday, (2010) categorizes SK – and to a limited degree LK – as conventional GIS knowledge, as it follows usual GIS knowledge structures. He categorizes the rest of the contribution continuum from LK to PK as unconventional knowledge.

Although different researchers typically categorize SK as structured and objective, whereas LK and PK are categorized as unstructured and subjective (Deparday, 2010), in this thesis a distinction is made between the context of the contribution and the nature of the contributed data, because LK and PK might be facilitated to be both structured and even objective. Indeed through appropriate contribution mechanisms, unstructured data can be channeled to a structured data format.

2.3.3.1 Contribution nature

Although in the definition of VGI, the term volunteer is present, the contribution is not always completely volunteered and various levels of willingness and desire to contribute can be identified. Here we differentiate between four types of contributions: Volunteered, f-VGI, Private, and Unbeknownst. The first level of contribution is a fully volunteered work, in which everyone is free to contribute. The term f-VGI, coined by Seeger, (2008) refers to the second level of contribution willingness, when individuals are asked to participate, for example to collect local knowledge within a planning procedure. The third level of contribution is very usual in social media applications, when the recipients of the participation are meant to be only a limited group such as friends or relatives (Elwood, 2008b; Sieber, 2007), however sometimes the shared information will be available for everyone either due to intended software design or user’s mistake in appropriate adjusting of the used software (Deparday, 2010). The fourth level of contribution is indeed an unaware and maybe unwanted contribution, when contributors do not know that they are contributing (Elwood, 2008b; Tulloch, 2008) such as posting images on the web using GPS-enabled devices that result in (unwanted) geo-tagging of pictures. Dobson and Fisher, (2006) call it “geoslavery” and describe this phenomenon as a "new form of human bondage based on location control". The discussion of geosurveillance among geographers goes back to 1991 at the "GIS and Society" meeting in Washington (Crampton, 2003). The dualistic nature of geosurveillance has prompted considerable debate among geographers, however most debaters agreeing with Dobson and Fisher, (2006) that "benefits do not negate risks".

2.3.3.2 Contributor continuum

Although most GI volunteers are categorized as amateurs (Tapscott and Williams, 2008), both professionals and amateurs can be volunteers. Literature offers many polarized definitions of professional versus amateur GI contributors. Coleman et al., (2009) summarizes and divides the contributor’s continuum into five overlapping categories (Figure 2-6):

1. "Neophyte": someone with no formal background in a subject, but possessing the interest, time, and willingness to offer an opinion on a subject;
2. "Interested Amateur": someone who has "discovered" their interest in a subject, begun reading the background literature, consulted with other colleagues and experts about

specific issues, is experimenting with its application, and is gaining experience in appreciating the subject;

3. "Expert Amateur": someone who may know a great deal about a subject, practices it passionately on occasion, but still does not rely on it for a living;
4. "Expert Professional": someone who has studied and practices a subject, relies on that knowledge for a living, and may be sued if their products, opinions and/or recommendations are proven inadequate, incorrect or libelous; and
5. "Expert Authority": someone who has widely studied and long practiced a subject to the point where he or she is recognized to possess an established record of providing high-quality products and services and/or well-informed opinions and stands to lose that reputation and perhaps their livelihood if that credibility is lost even temporarily.

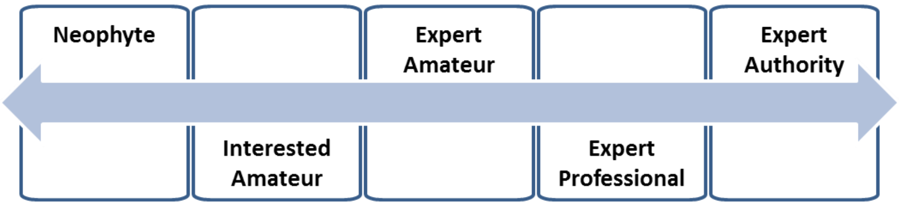

Figure 2-6: The continuum of VGI contributors (Coleman et al., 2009).

However, characterizing VGI contributors through these five divisions is rather a “simplistic” view of a multi-dimensional problem. For instance an “Expert Professional" might comprehend a company's mapping specifications and requirements as well as the restrictions of a provided GPS system, yet his knowledge of the characteristics or past history of a geographical feature may be restricted. A "Neophyte" contributor on the other hand may know little to absolutely nothing about positioning technologies or procedures yet is quite knowledgeable about mapped features in the area. Indeed, one maybe knowledgeable and expert in one area but rather a neophyte in another area of expertise. Goodchild, (2008) believes that “the old distinction between amateur and professional is quickly blurring in this arena, since few if any of the arguments that built and sustained the traditional system of map production are now viable”. Furthermore, he describes “a growing willingness of amateurs to be involved in the mapping process”, and a growing recognition that “we are all experts in our own local communities”, hence the need for new models to describe involved people and volunteers.

Further complexities are revealed through more empirical studies. They show only a very limited number of contributors produce the majority of the content. For instance, in OSM, as a well-developed VGI platforms, more than 80% of GI content is produced by less than 10% of the contributors (Budhathoki et al., 2010). This is even more obvious in cities like London, which only 2% of the contributors are responsible for 80% of the GI content, and 10% of the contributors generate near 95% of the whole content (Hristova et al., 2013). The situation is the same for other platforms such as Wikipedia, which even less than 2% of the committed, registered contributors, do 75% of the work (Kittur et al., 2007; Ortega and Gonzalez Barahona, 2007). This raises the question, why some individuals contribute much larger amounts of geographic information? And what is their motivation? Even more intriguing is to see a multitude of individuals, with no coordination by an official organization, collaboratively create something without any apparent financial benefit. Goodchild, (2007b) emphasizes motivation as an important condition for benefiting from VGI as a serious source of GI. Researchers suggest that the underlying motives are not primary altruism, but there are more complex motivation factors involved (Elwood, 2008b; Tulloch, 2008).

2.3.3.3 Motivation nature

What drives individuals to contribute GI, manage GI technical infrastructure, or development norms and policies for communities, what hinders individuals from contribution, how motivations connect to various degrees of contribution, and why motivations alter as individuals participate in VGI; understanding individual's motivations and inspirations is essential to design VGI procedures that harvest a greater contribution (Elwood, 2008b; Flanagin and Metzger, 2008; Haklay and Weber, 2008). In addition, motivation has potential implications and effects relating to the overall value of the geographical information, the credibility and reliability of the source, and the protection of personal privacy (Budhathoki et al., 2010). To better understand why individuals contribute GI and their motivation sources, lessons may be drawn from other fields such as free or open source software (F/OSS) and Wikipedia communities, because content generation in VGI has many similarities with the collaborative generation of knowledge in Wikipedia and open source software (Budhathoki et al., 2008).

Coleman et al., (2009) consolidate and summarize a list of motivators to make constructive contributions from empirical studies in the field of F/OSS and Wikipedia: 1) Altruism, 2) Professional or Personal Interest, 3) Intellectual Stimulation, 4) Protection or enhancement of a personal investment, 5) Social Reward, 6) Enhanced Personal Reputation, 7) Outlet for creative and independent self-expression, and 8) Pride of Place. Different mixture of these motivators can be applied for different VGI applications. For instance, altruism, professional or personal interest, and perhaps social reward are all solid motivators for people participated in stating particular instances or degrees of disasters, either natural or man-made (Crutcher and Zook, 2009; Laituri and Kodrich, 2008; Pultar et al., 2009). Pride of place definitely plays a significant part in motivating people to make updates to roads and point-of-interests in OpenStreetMap (Coleman et al., 2009). There have been also important negative motivators listed for contributors that are not interested in providing objective or reliable information such as; 1) Mischief, 2) Agenda, 3) Malice and/or Criminal Intent. In this research we consider only positive motivations that can be categorized into intrinsic and extrinsic factors and arise mainly from personal, social, as well as technological context of individuals.

There is a direct link between motivations and personality types (Deci and Ryan, 1985; Wagner, 1999). In addition, one's inspirations to do particular things can be related to where one is positioned in the human needs hierarchy (Maslow, 1987). Social structures and interactions also influence human behavior directly or indirectly. Indeed, different behaviors such as status gain, self-presentation, cooperation, and even altruism are rooted in social fabrics and depend on how the motivations are identified in a particular society; for instance one might choose to gather wealth as a symbol of status, whereas in another society a similar status can be achieved by education, or a authority position (Walsh, 1992). Social aspects have been considered as substantial motivators behind contributions to online knowledge and Open Source Software (OSS) communities (Hertel et al., 2003; Kuznetsov, 2006). Finally as we consider in this research only technology enabled VGI, a person's motivation and capability to contribute to VGI is heavily influenced by the degree of one's access to technological tools and one's ability to utilize them.

Intrinsic motivations	Underlying concept
Unique ethos	Distinguishing ideals, values, sentiments, or guiding beliefs that are shared by the members of a volunteering community.
Learning	A volunteer gets an opportunity to learn from his own experiences as well as the experiences of other members of the community.
Personal enrichment	A volunteer seeks to increase his intellectual or spiritual resources, which is found in the accumulation of cherished and valued experiences resulting from the chosen pursuit.
Self-actualization	It comprises the development and application of one‘s talents, capacities, and potential.
Self-expression	A volunteer seeks opportunity to express one‘s skills, abilities and individuality.
Self-image	It is enhanced through the expression of unique skills, abilities and knowledge.
Fun	An individual volunteer for hedonic gains that he derives from the pleasure of creation. Self-gratification or the satisfaction of one's own desires pertains to depths of satisfaction that may be at once fun, but can also be profound and fulfilling.
Recreation	It is the process of forming a new or creating one's self again; that is, volunteers retain a sense of renewal, regeneration, or reinvigoration through their participation in volunteerism.
Instrumentality	An individual volunteers if he believes that his contribution is crucial to accomplish the goal of the project.
Self-efficacy	A volunteer contributes if he perceives himself as having the knowledge and skills to meet the expectation of others in the team.
Meeting own need	When an existing product/service does not meet his own needs, an individual joins a voluntary community to collectively develop the product/service.
Freedom to express	An individual participates in voluntary activities as he has freedom to choose tasks and exercise his creativity.
Altruism	Volunteered action is directed by altruistic reasons.

Table 2-4: Potential intrinsic motivations for VGI (Budhathoki et al., 2010).

Budhathoki et al., (2010) have analyzed literature on volunteerism sociology, leisure studies, and social production of knowledge and open source software development communities and used them to provide a comprehensive list of potential intrinsic and extrinsic motivational factors for VGI (Table 2-4 and Table 2-5). Across societies the importance of intrinsic and extrinsic motivators is different. For example, open source software contributors in North American have a higher intrinsic motivation compared to Chinese and Indian (Subramanyam and Xia, 2008). Volunteerism is a fundamental principal in VGI (Elwood, 2008b; Goodchild, 2007b), however understanding volunteers‘ social and psychological constructs are essential to recognize what drives them to volunteerism. Clary et al., (1998) define volunteerism as an intended contribution where volunteers: a) often actively seek out opportunities to help others; b) may deliberate for considerable amounts of time about whether to volunteer, the extent of their involvement, and the degree to which particular activities fit with their own personal needs; and c) may make a commitment to an ongoing helping relationship that may extend over a considerable period of time and that may entail considerable personal costs of time, energy, and opportunity. Acts of volunteerism that seem rather similar and comparable on the surface may show significantly different various underlying motivational procedures (Clary et al., 1998).

Extrinsic motivations	Underlying concept
Career	An individual uses the voluntary work as a platform to signal his skills for career opportunities such as future jobs, a share in commercial companies or future access to the venture capital market.
Strengthen social relations	An individual volunteers to strengthen his social relations; participation in volunteerism depends on the reaction of his significant others.
Project goal	A volunteer carefully analyzes the goal of the project and its likelihood of attainment before participating in the activity.
Community	This pertains to efforts on behalf of the participants of a volunteering community to ensure that the community is maintained, continues to develop, and remains a cohesive unit.
Identity	By joining a group, an individual develops his identity with the chosen pursuit and is inclined to use this to identify himself. He also behaves according to the norms of the group.
Reputation	A volunteer contributes to enhance his reputation and continuously seeks recognition from his peers.
Monetary return	An individual participates in volunteering activities seeking a direct monetary benefit.
Reciprocity	An individual volunteers if he believes that others will reciprocate and will not exploit his contribution.
System trust	The volunteer‘s contribution depends on his belief about the reliability of the underlying technical infrastructure.
Networking	An individual participates in voluntary activities to network with other members of the community. The denser the network one has, the greater the contributions he makes.
Socio-political	An individual participates in volunteerism to meet her socio-political motives.

Table 2-5: Potential extrinsic motivations for VGI (Budhathoki et al., 2010).

Indeed the first step to have a successful VGI project is to comprehensively investigate the personal, social, as well as technological context of the individuals that are going to be involved in the project. Based on this analysis, contributors should be categorized to groups with quite similar contexts, and for each category a set of adequate intrinsic and extrinsic motivators (Table 2-4 and Table 2-5) should be carefully selected to do a specific task. Because potential VGI volunteers choose normally those tasks within a project – if there are different task choices – that best match their motivations (Houle et al., 2005). Deep understanding of the context of a VGI project such as the nature and continuum of contributors, contributions and motivations, is necessary to design appropriate contribution mechanisms that can best serve volunteers’ participation. In the next sub-section these different mechanisms are discussed.

2.3.4 Mechanisms

The mechanisms of VGI framework facilitate contributors with appropriate tools to produce GI content as well as address common problems faced in the production of content such as conflict, congestion, overuse, and quality. The mechanisms address actually the process of contributing GI such as; how people interact and cooperate, what are norms and rules are in place, what the data features are and how they are captured, what are the supporting processes and structures, and which evaluation mechanisms are possible. Careful analysis and evaluation of these mechanisms are necessary to understand strategies and techniques for gathering contributions and successfully implementing as well as executing VGI initiatives.

2.3.4.1 Contributory mechanisms

According to Olson, (1965) “unless the number of individuals in a group is quite small, or unless there is coercion or some other special device to make individuals act in their common interest”, even if there is a perfect consensus in a group and even with constituting individuals that are rational and interested in their common good, “no collective good may ever be obtained without some group coordination or organization”. Furthermore, the larger a group of individuals is, the more agreement, organization and coordination it will need to be able to reach a collective good. Olson, (1965) finds the differences between an organized and an unorganized group similar to the difference between a disciplined, coordinated army and an undisciplined, leaderless mob going to fight. Strong boundaries may encourage a healthy communication, yet they may also restrict the engagement level of the contributors.

Contributory mechanisms are informal and formal rules, norms or structures that make participants aware of set of dos and don'ts and define contribution boundaries and constraints. A boundary may be as simple as a limit for the number of characters in a tweet or any infrastructural limitation. These structures, rules, norms and boundaries are important to have a better quality of information and can increase the rate of contributions as well. Norms and rules-in-use are typically shared normative understandings that make participants familiar with what they must, must not, or could do in specific situations and circumstances (Hess and Ostrom, 2007) such as restraining from uploading GI materials that have commercial copyrights. A thriving online community normally defines these norms and standards itself, which gives a sense of ownership and freedom to the members of the community that is typically missing in official organizations and companies (Budhathoki et al., 2010).

When the community members start following the norms, they gradually become structures. Structures indeed evolve as an outcome of the interactions between the individuals and the technology and ultimately constrict the participants (members and non-members) as these structures determine what a participant can do or should refrain from doing (DeSanctis and Poole, 1994). OSM for instance provides some specific structures through its Application Programming Interfaces (APIs), which have been developed to let users import, export, add, edit, and tag data (Haklay and Weber, 2008). Although, structures and frameworks are flexible, thus can be altered when members of the community feel needed and necessary, they are crucial in determining the type and amount of information, that contributors and users are able to contribute, use or have access. Suggested OpenStreetMap tags are of this sort of flexible structures, which has been created to ease the maintenance of tags and to allow more enriched semantics analysis of tags. There is a trade of between freedom of members to do what they want and having strong structures; however structures may be designed in a way that give maximum freedom and guidance to the contributors and at the same time indirectly facilitate having a structured and clear data gathered.

2.3.4.2 Data capturing

Different capturing means, methods and approaches of geographic information have essential influence on different characteristic of the generated data such as format, quality and the accuracy of the data, because they directly influence the nature of the contributed data. Some of the most common approaches used by amateurs and professionals to produce spatial data consist of; Global Positioning System (GPS), Locating through the use of WiFi, Cell or IP loggers, geocoding, drawing on computer or paper-based maps. GI data are contributed through diverse devices such as smart phones, camera, tablets, PCs and so on.

GPS as one of the main tools to generate GI data is an evolved version of expensive professional GI devices that is now available as standalone or integrated in diverse consumer or professional devices such as phones or cars. Locating through the use of WiFi, Cell or IP loggers methods are detailed in (Turner, 2006). These methods are based on the fact that the internet network can give approximate locations by comparing the Internet Protocol (IP) of an address against a database of associating IPs and locations or by using the triangulation of signals originating from wireless base stations or a mobile phone network. These methods detailed in (Turner, 2006) and GPS are increasingly more and more combined with each other to locate diverse mobile devices such as cell phones.

Geocoding refers to the process of enriching the description of a location – typically an address or a place name – with geographic information through associating it to a geographic location or locating it on a map. Geoparsing is a technique related to geocoding (Scharl, 2007), which consists of scanning unstructured documents, extracting all geographical references, and assigning geographic identifiers such as geographic coordinates to them or locating them on a map using geocoding techniques. “30 miles south of Manhattan, NY” is e.g. a textual geographical reference which can be a written or an audio content. Geocoding handles normally structured location references, whereas geoparsing should analyze rather unstructured and ambiguous location descriptions.

One of the most widespread methods in web mapping is drawing on a computer-based map, where individuals can add content to an overlay layer on a map that provides the base data such as a satellite image. It can be as simple as adding a point on the map and attaching a comment to it or in cases such as OpenStreetMap one may draw complex lines and polygons and add different semantic information to them. Drawing on paper-based maps to collect VGI content is still practiced in some cases such as OpenStreetMap walking papers project that enables contributors to print paper maps to draw on in the field and scan them later on to automatically create digital information.

2.3.4.3 Spatial data features

Spatial data features of contributed GI are heavily influenced and defined by the data capturing approach and the type of features that the applications support. Some applications allow just insertion of points whereas others enable contributors with basic geo-processing capabilities such as measuring, merging and union, and also the main drawing tools for sketching the maps such as adding and modifying lines as well as polygons. The richness of the available spatial data features have essential effects on the richness of the collected GI, however they may cause different challenges and difficulties when analyzing and interpreting them.

2.3.4.4 Attributive data features

Attributive data features relate to the nature of the textual and semantic data. In GIS, almost all geospatial data contain some attributive data and information about the features; hence, for evaluating the accuracy of GI data attributive accuracy should be also evaluated (Chrisman and McGranaghan, 1990; Lo and Yeung, 2007). The attributive accuracy is “defined as the closeness of attribute values to their true value” (Chrisman and McGranaghan, 1990). The attributes – unlike locations – dose change over time and depending on the nature of the data must be analyzed differently (Chrisman and McGranaghan, 1990). An important remedy for quality problems associated with GI attributes is using data standards and structures (Lo and Yeung, 2007) such as OSM tagging standards. However, in the context of VGI, in order to give sufficient freedom and flexibility to individuals to express and share their personal and local realities and facts, alphanumeric information generated through VGI capturing methods often have a much more flexible structure.

The lack of structure or metadata as a key issue of VGI semantic data comes from the fact that the ever-increasing amount of data is generated by a multitude of authors, not a group of professionals that have a structured description of the data through attributes or metadata. An efficient strategy to cope with this issue is to benefit from the power of the contributors themselves to categorize and classify the data and involve the authors and users directly in the development and creation of metadata and categories. This approach of creating metadata is coined as folksonomy mirroring the "folks" origin of the taxonomy. Smith, (2008) presents a comprehensive review of different tagging (folksonomy) techniques such as leaving total freedom to the contributors to tag, suggesting tags dynamically based on the used tags, or enforcing a fixed tagging structure to which the individuals have to adhere such as predefined categories and tagging structures. As discussed, an adequate balance between having a structure and freedom is required, which varies depending on the VGI project.

Depending on the aim of the data collection and the varying approaches for dealing with the data in different VGI projects, the collected data may be structured or unstructured and at the same time objective or subjective. Structured and organized data collection works based on a specific plan and entails certain information about the information content and its possible units that are to be collected. Indeed to have a structured data, the various features of the information that are to be gathered are decided in advance and may include using especial data collection tools that are likewise structured in nature. However when it comes to unstructured data collection, its basics are essentially different from structured data gathering. In the case of unstructured data collection, individuals have the freedom – like free texting – to record and gather what they feel is appropriate and relevant about the information.

Structured attributive GI data refer to attributes that are linked to spatial features and “conform to a range of values on nominal, ordinal, interval or ratios scales” (Deparday, 2010). Unstructured attributive GI data refers to free-text comments and opinions that are linked to different spatial features. VGI data have normally a combination of both structured and unstructured data that are associated to spatial features. For instance a VGI initiative that gathers information about accidents in a region can allows the contributors to add unstructured free text comments or pictures and at the same time choose between categories available for the type of the accidents.

Any data collection process consists of two main components; the subject and the object. The subject refers to the data collector – here contributors – whereas the object refers to the observed information, process or activity, about which GI information is going to be collected. Subjective information entails one's own immediate experiences whereas in collecting objective information the individual observer is apart from the thing that is observed. Tulloch, (2008) differentiate objective and subjective volunteered geographic information as purported facts versus offered opinions.

However facts may be presented differently by different individuals; which makes semantic analysis of the attributive data difficult. Hence, providing a standard terminology for contributors and users of VGI will prevent later difficulties in utilizing the collected information. Furthermore, it may increase the quality of the data significantly.

2.3.4.5 Action and interaction

The action and interaction mechanisms of the VGI collection process relate to the possible level of contributors interaction and involvement. The action and interaction levels are a central concept for PPGIS frameworks, which has been represented diversely such as through the public participation ladder (Arnstein, 1969; Schlossberg and Shuford, 2005). An extensive listing of different involvement methods is presented in (Rowe and Frewer, 2005) that has been further characterized and detailed by a typology in (Turkucu and Roche, 2008). Rowe and Frewer, (2005) differentiate between three types of public engagements based on the flow of information between participants and sponsors (the party commissioning the engagement initiative); namely public communication, public consultation, and public participation.

The information flow in public communication is one-way from sponsor to the public and there is no involvement of the public because the public input and feedback is not sought. This first level of information flow is not relevant in the context of VGI. In public consultation, information flows from participants of the public to the sponsors of the initiative, following a process launched by the sponsor. At this level also there is no formal dialogue between participants and the sponsors. This level of interaction is also rare in today's VGI initiatives. In public participation, information is exchanged between participants and sponsors. Although Rowe and Frewer, (2005) consider the interactions mainly between members of the public and the sponsors, most web 2.0 applications nowadays make it possible for the facilitators to have a full interaction and communication with the members of the public. This full interaction is even possible between the members of the public themselves (Hall and Leahy, 2011). This interaction level is crucial for the VGI efforts, because contributors and users are often becoming synonymous in the context of VGI (Goodchild, 2008b).

Deparday, (2010) considers this as an n-way flow, where "users can also cooperate in the data collection by commenting or editing each other's contribution". So we have three degrees of interaction for VGI initiatives, 1-way from the contributors to the facilitator, 2-way between facilitators and contributors, and n-way from contributors to contributors. In 2- and n-way flow of information contributors have the possibility to see the outcomes of their work integrated with the contribution of other individuals. Sometimes in the final level of involvement, a facilitator does not exist but a community of contributors that plays the role of the facilitator. These three kinds of involvement are completely different both structurally with different enabling mechanisms as well as their objectives.

2.3.4.6 Evaluation mechanisms

One of the most important concerns about VGI is data quality and its credibility (Flanagin and Metzger, 2008). This concern is mainly a result of volunteer's demographics, motivations and inspirations, and their abilities. As data are gathered for a better understanding of the phenomenon under investigation, data accuracy and reliability problems have to be thoroughly reviewed and evaluated. To deal with the issues of data credibility, accuracy and quality, not only the gathered data should be evaluated, appropriate mechanisms should be in place to help and facilitate volunteers with gathering geospatial information. For instance facilitators may provide volunteers with guidelines and standards that can be used to understand the best ways to maintain data quality at an optimal level (Dickinson et al., 2010; Newman et al., 2010). Indeed, if volunteers are guided appropriately, the collected information tends to be of a higher quality. The contributed information by volunteers can additionally be cross-verified with the information gathered by researchers, or with a dataset from authoritative organizations, in a way that both sources of data complement each other (Clark and Aide, 2011). An additional approach for controlling information quality – like Wikipedia – is to review volunteer contributions extensively by several moderators who have the authority to correct and fix possible mistakes, before making the information public (Goodchild and Li, 2012).

Furthermore, contributors not only can collaborate in contributing data and utilizing the resulting datasets, importantly they can correct contributions of others, which consequently can significantly improve the quality of gathered data. This is in accordance with the Linus' Law from the computer science world; "Given enough eyeballs, all bugs are shallow", which means if sufficient individuals with adequate abilities look at an issue, all obstacles can be conquered (Haklay et al., 2010; Raymond, 1999). In case of VGI, the Linus' Law can represent the number of contributed points, lines, polygons, or attributes by volunteers for a specific object on the map. Hence, by increasing the number of contributions over time, data credibility, accuracy and quality increases (Haklay et al., 2010). Mummidi and Krumm, (2008) propose when VGI is aggregated from several sources, the quality of the information increases. Goodchild, (2008b) presents also that user generated content can potentially offer an effective system for correcting and modifying errors. However, empirical studies on OSM show that the relationship between number of contributions and quality is not linear and in most cases having five or more contributions for an individual object on the map can lead to a high quality of data representation (Haklay et al., 2010).

2.3.5 Utilization of VGI information

The utilization of the VGI framework consists of two phases; a) data aggregation, filtering, and quality check, and b) information synthesis and developing innovative services. Quality control and aggregation of the gathered data from different data sources are of vital importance for almost any VGI initiative, without which the developed services cannot successfully provide a solid basis for making decisions. Numerous studies have focused on the utilization of the VGI and analyzed different aspects of the gathered data such as – among others – trust and credibility (Bishr and Mantelas, 2008; Flanagin and Metzger, 2008); quality and coverage (Clark and Aide, 2011; Haklay et al., 2010); privacy and control (Harvey, 2007); access and empowerment (Tulloch, 2008); and overall effect on social and political procedures (Elwood, 2008b).

The results of individuals' contributions constitute indeed a public good that lies at a different point of the public-private continuum. Public goods are distinguished by two features: indivisibility, indicating that a person 's consumption of the good does not lower the quantity available to others; and non-exclusiveness, indicating that it is difficult or even impossible to exclude people from benefiting from the public good (Kollock, 1999). For instance, while OSM and Google Map may appear to be similar – at the surface – from the public good point of view, their underlying copyright laws and regulations are noticeably different (Budhathoki et al., 2010). These characteristics of the

VGI data that are inherited from VGI being a public good literally set the sky as the limit for its utilization. The proof of this fact can be seen in numerous services, utilizations and researches that have been developed from diverse VGI datasets, which are increased further daily.

2.3.5.1 Data aggregation, filtering, and quality check

Contributions in VGI "are the aggregate repository of user-contributed geo-referenced information. Such a repository may contain different types of information (e.g., points, lines, polygons, images, pictures, or text) depending on the goals of a specific VGI project" (Budhathoki et al., 2010). Individuals may also make other contributions such as raising issues, commenting on raised issues by others, and engaging in conversations with other participants of the community (Budhathoki et al., 2010). Indeed, there is an overload of voluminous information which is too large to be browsed and edited manually. Furthermore, the contribution procedure typically has a limited structure. Therefore, the information overload is intensified due to the heterogeneity and diversity of the contributions. Contributed spatial data are often cluttered and overlapped and the free-texts with mainly different semantic information and language standards make the whole information utilization phase more tedious.

Although, in some VGI initiatives, some of these challenges can be reduced by enforcing more structure such as categorizing data on the contribution procedure, these structures may reduce the maximum contribution potential while not being able to fully eliminate the issues. Hence, different approaches such as aggregation and filtering methods are necessary to reduce the overload of information based on data quality. Indeed, through automatic and manual approaches erroneous information should be discarded, duplicate data should be aggregated, and meaningful semantic information from the data should be extracted to maintain a specific level of quality in the datasets. VGI errors have been revealed not to be arbitrary, and happen across spatial as well as thematic domains generating data that need to be assessed and evaluated before operational usage (Girres and Touya, 2010; Haklay, 2010; Zielstra and Zipf, 2010).

In order to synthesis information and reveal structures and patterns in datasets, different grouping and aggregating methods and complex techniques for knowledge discovery and data mining are used. Although, the use of automatic processing techniques with VGI data aims to simplify the data, generate categorizes and summary information to reduce the quantity of the presented data, the employed approaches must allow the individuals to have access to the original information. This is crucial to prevent black box situations and maintain a transparent and clear procedure. This is more important in political and social contexts, hence, except in apparent incorrect and abusive data entries, the data should not be suppressed (Deparday, 2010). Slocum et al., (2008) presents several traditional cartographic methods that can be used to reduce the overload of spatial information in making maps. Multitude of natural language processing algorithms can be also used to categorize and summarize free form texts connected to the spatial components of the data.

Another method to reduce the information overload is to filter out not relevant information to the users and allow the individuals to find and retrieve relevant subsets of information (Roberts, 2008). This is typically accomplished using filtering systems and dynamic querying methods that allow the users to reduce the information to a part of information. The subset of information is specified according to different dimensions of the dataset such as location, contributor's name, time, semantic content, and the generated structures via data aggregation like categories, classifications and clusters. A combination of different Boolean operations may be used to filter out according to several dimensions simultaneously. After aggregating the collected data into databases and increasing its quality, the databases can be made publicly available as raw data or through graphs, charts, interactive maps or other innovative services.

2.3.5.2 Information synthesis and innovative services

As discussed, the complex multifaceted, heterogeneous, qualitative and subjective nature of VGI makes its utilization a challenging task. Nevertheless, a data with such characteristics and features constitute a richness that should be exploited and changed into useful information. Moreover, this richness is become even greater when considering the dataset of all of the contributions together as it can disclose relationships, connections, patterns and insights that cannot be achieved when analyzing contributions separately.

By utilizing VGI, researchers are engaging volunteers for initiatives such as wildlife preservation, mapping invasive species, monitoring radiation degrees (Silvertown, 2009). The Christmas Bird Count project initiated by the National Audubon Society (established in 1900) to monitor changes in the population of birds is recognized as one of the pioneering examples in VGI. Aside from many examples related to biodiversity monitoring, VGI has been also used in remote sensing. For instance, Clark and Aide, (2011) used the Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT) to enhance the accuracy of reference samples via cross-verification of human as well as professional interpretations and analysis. This combination helped to successfully address problems associating with land change. As another example which illustrates the benefits of engaging volunteers with remote sensing imagery, we can refer to Geo-Wiki Project[5] that uses volunteers to improve the classification quality of remotely sensed global land cover maps.

Emergency management, disaster response and land management are other ways that communities and governments are using VGI (Haklay et al., 2014). As an example, during the Haiti earthquake in 2012, with the helped of Ushahidi[6] volunteers contributed vital information such as trapped and injured people, damaged facilities and infrastructure, and closed or open roads. Moreover, VGI is used in enhancing the temporal accuracy of outdated information and datasets (Sui et al., 2013). As an excellent example of this we can refer to OpenStreetMap (OSM). The Canada-OSM Synergy Project which has been launched by Natural Resources Canada (NRCan), GIS data such as road networks are uploaded to OSM in order to be modified and kept up-to-date by citizens (Haklay et al., 2014). There are other mobile applications which have been created by municipalities such as the City of Calgary to enable residents in reporting specific (geo) concerns on bike paths, streetlights, snow and ice, water and sewer services, traffic signs, street and park lights, and road maintenance. City authorities utilize these information to reduce costs and expenses, and deploy resources where required.

Aside from previously discussed initiatives, the involvement of volunteers can be highly beneficial for diverse other fields such as in scientific research. Silvertown, (2009) notes "research

[5] Geo-Wiki Project: Improving the quality of global land cover maps with volunteers – geowiki.org
[6] www.ushahidi.com

funders such as the National Science Foundation in United States and the Natural Environment Research Council in United Kingdom now impose upon every grant holder to undertake project-related science outreach". He suggests that this inclusion of volunteers in research can be an exceptional opportunity to spread out knowledge and educate the populace at the very same time. Bonney et al., (2009) showed how their study raised scientific understanding amongst their participants on ornithology, particularly because it related to bird watching. In addition, evaluating more than 230 projects by the UK Environmental Observation Framework (UK-EOF), it was concluded that volunteers provide information with a "high value to research, policy, and practice" (Tweddle et al., 2012).

3 EVENT NOTION

Geographic Information Systems initially modelled geographical features and their relations independent of time. The main reason for this assumption was that most of geographical features maintain their identity as well as their location for relatively long periods of time. Due to the perseverance of these essential and basic properties, representation of changes in time was not a preliminary consideration for Geographical Information Systems. Furthermore, early spatial information-collection approaches concentrated only on capturing and recording such fundamental properties like identity and location. In addition, very high expenses usually made it impossible to repeat the capturing with a frequency that could possibly manage appealing change analysis and evaluation. Towards end of 80s and beginning of 90s, GIS society started to consider time in GIS and address the dynamic aspect of geographical features (Armstrong, 1988; Langran, 1992). This enhancement, for the first time provided the possibility to record the history of objects with their attributes which could be used to predict possible future changes. The main emphasis nevertheless continued to be geographical features and temporal dimension was added as time stamps to keep track and evaluate different states of features. In this object change view – coined by Worboys, (2005) – objects have a unique identifier which sustains, and changes may happen to both spatial and non-spatial features.

In this chapter, we start by distinguishing two views on the objects and processes and their relationships. The conventional object-oriented view is clearly reflected in the ontologies that have actually dominated western thoughts at the very least since the Aristotle time (Rescher, 2008). The process-oriented view has consistently appeared in the history of philosophy, as far back as Heraclitus, as well as more recent philosophers such as Bergson, James, and Whitehead (Galton and Mizoguchi, 2009; Rescher, 2008). In the object-centered view, matter and objects are considered prior to processes and events. The 'object-priority' view claims that essentially matter and objects are all that exists in the world; hence the existence of events and processes is entirely because of the distribution of matter in time and space. In the 'process-priority' view "processes and events are prior to matter and objects". This view is also normally presented as an ontological claim that only processes and events exist in the world and objects and matter are constructed or emergent from processes and events (Galton and Mizoguchi, 2009). "It is sometimes said that what we commonly call an object is in fact an event or process, although it is hard to assert this using ordinary vocabulary without courting gross categorical confusion" (Galton and Mizoguchi, 2009).

The object-oriented method has been presented and applied to spatial information modeling in (Worboys et al., 1990). This view sees the world as a collection of classified and identified objects, with specific properties and relationships. Nevertheless, there is an even increasing amount of work revealing that in numerous cases, the dynamic facets of geographical phenomena are vital to create useful explanatory, informative and predictive models. Hägerstrand, (1970) highlights the importance of time in human activities to assess the dynamic behavior of people in space, especially the motion of individuals in space and time. Miller, (2003) and Yuan, (2001) have exhibited this fact in their work on "transportation and urban analysis", and on analysis of physical phenomena, such as storms. Different researchers such as Mark, (1998) and Miller, (2003) have promoted the work of Hägerstrand, (1970) under the principal of geo-spatial lifelines. However, Hornsby and Egenhofer,

(2000) deal with the object change view through the concept of identity-based change. In the past nearly three decades, since the initial consideration of time aspect in GIS, various spatio-temporal information models have been suggested to handle storage and management of data (Abraham and Roddick, 1999; Langran and Chrisman, 1988; Langran, 1992; Worboys, 2005). Both views will be discussed in more detail in the following section.

3.1 Three phases of spatio-temporal modeling

As shown in (Peuquet, 2005) and (Worboys, 2005), the advancement of spatio-temporal modeling has actually undergone three evolution phases: the initial stage is snapshot modelling, the second phase is object-based modelling and the third stage is event-based modelling.

3.1.1 Snapshot model

Indeed, the most common method for spatio-temporal modelling of the world has been viewing the world as sequences of temporal snapshots of objects with different spatial configurations and setups (Worboys, 2005). These systems however enable representation of only a solitary state of information about the considered object state (Worboys, 2005). Typically the most current state is considered as the most interesting one; therefore database updates try to maintain it as current as feasible. Although, it is possible to represent past or future states of objects, merely an instant in time can be represented, therefore it is not possible to compare states at different times (Worboys, 2005). Indeed, a temporal snapshot is a depiction of an event's state in a specific domain at a certain moment in time. Temporal series are a compilation of temporal snapshots, typically from the very same spatial area, indexed with a temporal variable. The snapshots can be considered as sampling a dynamic phenomenon at specific temporal sequences. For instance a temporal sequence can contain some temporal snapshots of a residential area, referenced to different years; 1950, 1975, 1995, …, and 2015. As time passes several changes may take place, such as construction of new buildings. This instance shows clearly the significance of reflecting the real-world changes to the databases.

The general forms of temporal queries to such systems that represent changes implicitly through a series of static snapshots is "What was the state of this object at that time?" or "At what time did this object have that state." When it comes to spatio-temporal information queries, the query ends up being "Where was this object at this time?" and its converse "At what time was the object at this location?" Considerable amount of work can be found in the literature on temporal and spatio-temporal models and query languages (Abraham and Roddick, 1999; Snodgrass, 1995; Worboys, 1994). In theory, these systems can be used to fully represent changes of geographical features and attributes, nonetheless, through extensive saving of all variations of information this can be achieved. Furthermore, there are three main downsides when modeling changes with snapshots (Langran, 1992):

- Relatively expensive computations and calculations are needed to detect and identify changes between snapshots, which are due to the fact that both snapshots need to be compared extensively.
- Developing or imposing rules for internal reasoning is challenging, since there is no understanding of the restrictions upon on the temporal structure.

- No matter what is the size of changes, a full snapshot is produced at each time sequence. This replicates all unmodified information, which leads to storing huge amount of redundant information.

Establishing and maintaining the identity of objects over time is another challenge for the object change view. This question normally emerges as when a change is so substantial that a particular object is not any more the very same object (Beard, 2006).

3.1.2 Object-based model

The snapshot model provides no system for explicitly representing the time and the incident of events such as construction of a new building. The advancement of spatio-temporal modeling in object-based models follows the sequential updating idea based on snapshot models; however it just keeps changed elements. This approach has been developed by Hornsby and Egenhofer, (2000). The main difference of this model with the snapshot model is in shifting the focus from temporal sequences of static images composed of objects, their features and relationships, to change itself, which happens to objects, attributes and relationships (Worboys, 2005).

The difference between first stage (snapshots) and second stage (object change) model can be further described with reference to Figure 3-1. The figure shows different sequential snapshots representing the development of a residential area from 1950 to 2015. By adding the following list of temporally referenced changes, an object based model is represented.

- 1950–1975: A road is added. A school is built. A tree is planted.
- 1975-1995: A house is burnt down. Two new houses are built.
- 2000-2015: Two buildings are built.

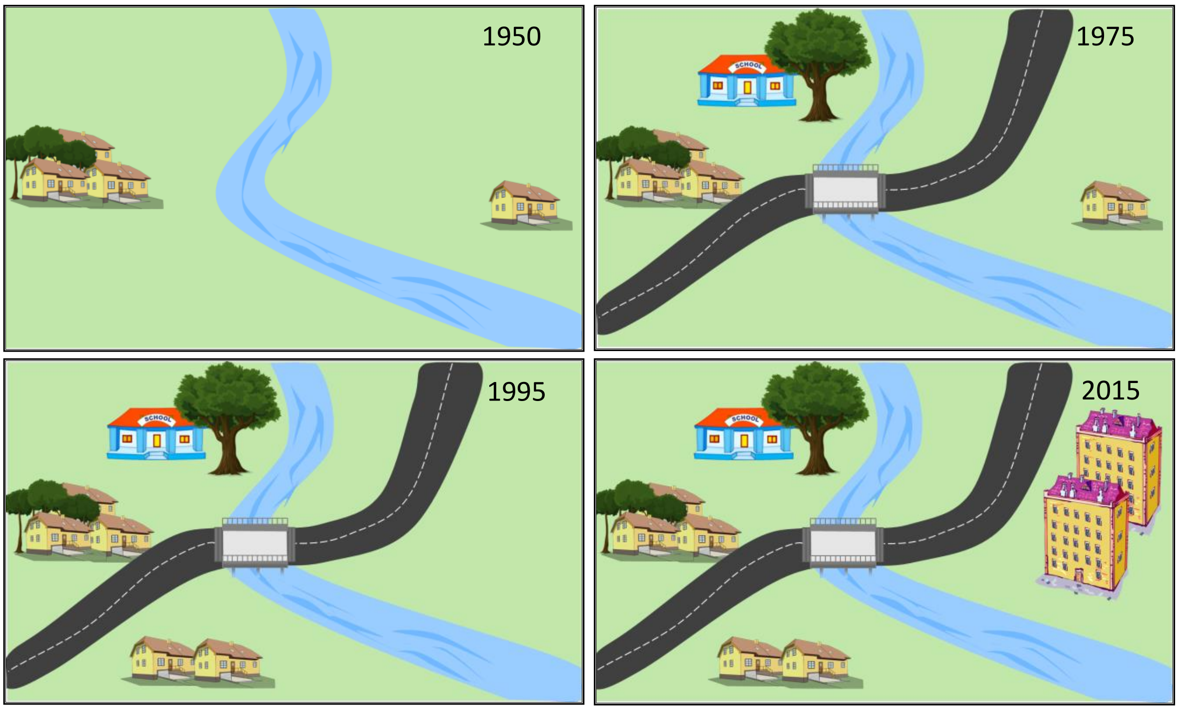

Figure 3-1: Evolution of a residential area.

Although, a world model based on the evolution of objects in time, which preserve their identification yet altering spatial and none spatial characteristics, appears to be natural, problems emerge particularly in connection to the continuity of identity over time (Worboys, 2005), due to the ontology changes of physical objects over time (Heller, 2008). Furthermore, problems may arise using this model when facing hybrid changes, in which both spatial and non-spatial attributes change and the object moves such as an army, a wildfire or a spread of an infectious disease (Worboys, 2005). To further develop the model, a collection of change "primitives," such as creation, destruction, appearance, disappearance, transmission, fission, and fusion were developed. More intricate changes are created from combining these primitives (Hornsby and Egenhofer, 2000; Worboys, 2005).

These terms in their own turn are indeed events which happen to objects. In order to explain complex changes, modeling events is essential. The third stage of spatio-temporal modeling (event-based modeling) is developed to achieve this goal.

3.1.3 Event-based model

Both snapshot and object-based models can be considered as an extension to conventional vector and raster representations. They excel at executing location- and feature-based enquiries, yet not appropriate for analyzing and evaluating temporal relationships of events and their patterns (Peuquet and Duan, 1995). Hence, due to such restrictions in these models that merely include time stamps for managing versions and changes to the state of geographical attributes or locations, many researchers have suggested event-based models as an alternative solution (Claramunt and Thériault, 1995; Peuquet and Duan, 1995; Worboys and Hornsby, 2004; Worboys, 2005). Although early calls to maintain and preserve records of events and processes in order to understand dynamic behaviors go back to late 80s (Chrisman, 1998), “the realization of the event view owes much to new technologies that are now able to deliver a wide range and volume of spatiotemporal information” (Beard, 2006). Big repositories of information with high temporal resolution are created by environmental monitoring and sensor data streams to analysis occurring changes. Many temporal data streams with fine resolutions help in understanding how processes are working, which will be a basis for exploring cause and effect relations. In addition, considering that data is increasingly available, renewed impetus to develop tools and models for managing events and processes are needed (Beard, 2006).

In event-based models, change is the main concept that is modelled and change units are the primary items for analysis and evaluation. In this approach the spatial dimension is dominated by the time dimension, because the sequence of events in time is essential (Beard, 2006). As the primary units of analysis and the dimensions are essentially different, new methods are required for modeling changes in objects, in ways that objects participate in events and relationships between events are modeled. A truly event-oriented approach should enable us to proceed from simple snapshot queries such as "What happened at this place at this time?" to a much richer language which entails the interaction between objects and events, and event-event relationships.

In an object change view usually changes in the attributes of an object are recorded (Beard, 2006). For instance, a house that is painted from white to yellow undergoes a non-spatial alteration to its color attribute. The main object here is the house that its new color is recorded and perhaps its previous color would be kept. In an event view, 'painting the house' is the unit to be recorded together with its particular attributes such as start time, period, and possibly the method of accomplishment. In an event view the emphasis relocates from the change to a certain house to an analysis of the changed objects themselves, such as analyzing painting events via for instance comparing their seasonality and durations. An example query can highlight the differences in the perspectives better. Through GIS, an individual may query all houses that have been painted over a specific year, obtaining a map that shows the geographical position of these houses. In this case, we emphasize on geographic objects with a specific kind of change. Whereas, in an event-oriented model, all 'house painting' events over a specific year are searched. Although, the set of house painting events would have the same distribution, different sets of characteristics can be queried, analyzed and evaluated (Worboys, 2005). The main objective of event-based approaches is to model changes explicitly to facilitate the analysis and evaluation of changes, their patterns or occurrence through time (Worboys, 2005). Galton, (2004) makes "the distinction between histories that are functions from a temporal domain to attribute values, or properties of objects, and chronicles that treat dynamic phenomena as collections of happenings".

Claramunt and Thériault, (1995) define events as things which occur. Particularly they explain that processes cause changes in the state of objects, these changes reveal the outcome of the process and create events. Peuquet, (1994) defines an event as indicator of changes in a place or an object. Peuquet and Duan, (1995) refer to an event as a way to represent spatiotemporal manifestation of processes. Worboys, (2005) and Worboys and Hornsby, (2004) define an event as a happening that should be differentiated from a thing or continuant. They comment that the main weakness of snapshot models is the lack of an explicit representation for events and changes. Furthermore, they suggest that events are necessary to record the mechanism of change.

From an ontological point of view, entities that exist in the world should be divided into continuant entities, that endure with time (such as tables, homes, and individuals) and occurrent entities, that take place or happen (such as talks, individual's lives, races) (Worboys, 2005). There is a distinction between a metropole, whose features are captured by demographics and surveys (each decade), and the procedures of city growth and development or decline, migration and movement, and development, that comprises the city in change and motion (Worboys, 2005). Grenon and Smith, (2004) designate temporal sequences of object's different configurations the SNAP ontology, and the "event/action/process" view, the SPAN ontology. Event based model belongs to SPAN.

The proposed Event-Based Spatiotemporal Data Model (ESTDM) by Peuquet and Duan, (1995) assign changes (temporally) to places within a pre-defined geographical location. While changes are explicit in this model, they are discreetly subservient to a temporal and a spatial information (i.e. bound to pixels) (Beard, 2006). ESTDM is like temporal map sets. It is a raster-based model, which utilizes a series of timestamped elements to represent an event; nevertheless, it saves changes in connection with a previous state instead of a snapshot of instances. Certain changes are related to a saved temporal area called a time stamp in an event list. These lists are related to a single thematic layer. For each single theme of interest a base map is needed as a preliminary snapshot. The ESTDM has exposed its abilities and performance in both spatial and temporal queries. The ESTDM labels itself as an event-based information model; however it does not deal with change as the main unit. The main distinction is that the emphasis is on a change related to a place (a pixel); hence in an event model instead of a “geographical feature or a location”, the change unit must be characterized (Beard, 2006).

Claramunt and Thériault, (1995) suggested an event-based method to model changes amongst a set of entities. Their model connects spatial entities and their temporal variations through logical intermediary tables of past, present and future events to ensure the description of complex succession, production, reproduction, and transmission processes. Nevertheless, they have left out specific sort of change therefore no systematic dealing with change have actually been carried out (Hornsby and Egenhofer, 2000).

Allen et al., (1995) proposed a generic model to explain the relationships among events and to explicitly represent casual links in a spatio-temporal GIS. In this model a small number of elements are presented through an extended Entity-Relationship formalism, consisting of objects and their states, events, agents, conditions, and relations, leaving out the causality between events and agents. Chen and Jiang, (2000) suggested an event-based method for modeling spatio-temporal procedures, where events and their causal relationships with states (event-event, event-state, and state-state) were stored. The model however is over specified on procedures in land subdivision.

The literature has very well expressed the motivations for an event-based approach. Event-based modelling supports “representation of dynamic behaviors of geographical phenomena, hypothesis generation, scientific investigation of complex relationships, an ability to investigate causal linkages and associated entities with influences and underlying procedures” (Beard, 2006). Another practical usage of event-based models is utilizing available temporal information from monitoring and sensor networks. These data have remarkable value yet not completely utilized due to limitations in integrating and assimilating heterogeneous spatial and temporal information. However, most of the researches in this area are focused mainly on modeling events and relationships between them. Time information is included as an attribute for spatial entities or as an integral part of spatiotemporal objects. Furthermore, semantic information has actually not been much taken into consideration. Hence the main goal of this thesis is to present a conceptual model, which builds a high level geo knowledge-based system to manage real world processes such as the relationships amongst objects with each other, objects with events, events with events and the involving processes. This study uses a generic event-oriented perspective to implicitly represent causal relationships among different components of a Spatio-Temporal Geographic Information system. In this new perspective, the objects in space and time are considered merely as information elements of the events, which are connected to other event elements through internal or external processes. In the next sub section, first the notion of event is described.

3.2 Event Notion

Events play a prominent role in various research areas such as physics, philosophy, psychology, linguistics, literature, probability theory, artificial intelligence, and history. One can find many technically refined concepts of events and objects across these diverse disciplines; while many scientists look for a commensurable notion for events and objects across disciplines. There is a tension between the latitude of the notions and individualization of notions in each discipline; one tries to unify whereas the other tries to divide it.

Four types of notion in the taxonomy of events are as following (Casati and Varzi, 2008; Casati, 2005):

- Pre-theoretical, common-sense (CS) notion
- Philosophically refined (PR) notion: refinement/replacement of the CS-notion
- Scientifically refined (SR) notion: refinement/replacement of the CS-notion
- Psychological notion: the I-representation ('I' for 'internal) of the CS notion; explanation of cognitive aspects of CS-notion.

A very simple example that can show the differences between CS and PR/SR notions are material objects. In CS-notion, objects are three dimensional entities in space which continue to exist wholly through time during their existence (here, events are defined in 4D). Whereas, the refined PR and SR notions consider material objects as four-dimensional entities in space that exist only partly through time during their existence (physicians come to spatiotemporal 4D parts from Relativity Theory while philosophers conclude it from Change theory). It can be deduced that properties which are assigned to an object in revised notions (PR/SR) are the same with the properties that a CS-notion uses for an event (Balashov, 1999; Casati and Varzi, 2008). Finally, I-representation tries to explain/recognize certain linguistic or logical performances from the event. So implicitly represented entities in a statement can be deducted and justified.

Different scientists have very different ideas about events and a core notion (Bennett, 2002; Casati and Varzi, 2008; Lewis, 1986; Von Kutschera, 1993). Casati and Varzi, (2008) proclaim that a common core notion might be not possible, unless one uses a PR notion. PR notions have also some significant divergences, but they share an invariant common core of characteristic features; every event has some objects (e.g. participants) and involves in some relationships (e.g. causal relationships) with other events (Casati and Varzi, 2008). PR notion considers conceptual interactions between CS-notion with other notions.

Casati and Varzi, (2008) listed the most important conceptual interactions between common-sense notion and other notions as following:

1. How does the event notion interact with our understanding of causality?
2. How does it feature in causal explanations?
3. How does it interact with our concepts of time and space?
4. How does it interact with the notion of intentional action?
5. How does it interact with the notion of an object?
6. How does it interact with the concepts of identity and individuation?

A close study of these interactions aside from event notion is of vital importance in understanding and developing a very conceptual upper ontology for events.

3.3 What events are

Different scientists have studied spatiotemporal ontologies (Agarwal, 2005; Schuurman, 2006). Nevertheless, a lack exists in explicitly considering events as entities in geographic information systems (Galton and Worboys, 2005; Hornsby and Cole, 2007; Klippel et al., 2008, 2006; Liu et al., 2008; Mau et al., 2007; Worboys and Duckham, 2006; Worboys, 2005). The variety of event notions can be compared with the diversities in definitions about events. Furthermore, different researches does not agree on using semantic terms such as “events”, “process”, and “states”, however they emphasize their value for GIS (Weiser et al., 2012). Remarkable works in this area include formal ontological methods by Grenon and Smith, (2004), conceptual models by Worboys and Hornsby, (2004), and philosophical implications by Galton and Mizoguchi, (2009).

The philosophical differences between continuants and occurrences or incidents are a starting point to discuss processes and events in the official community of ontology. These differences have been summarized in the differences between SNAP and SPAN ontologies promoted initially by Barry Smith and his partners (Galton, 2006a). SNAP ontology supporters are continuants; “entities that have continuous existence and a capacity to endure . . . through time even while undergoing different sorts of changes”, while SPAN ontology supporters are occurrents; “processes, events, activities, changes”. A SNAP ontology consist of elements; “all continuants existing at some given instant of time”, so obviously the classification that SNAP recommends is the concept of a snapshot, a full picture of the universe at a moment. On the other hand, a SPAN ontology, extends a series of instants, particularly it includes entities that their nature contains such spans (Galton, 2006a).

What a snapshot reveals is exactly all that existed at the time it was taken. However, what we experience is a dynamic globe, not static and fixed. In a snapshot we can presume the walking; in a real experience, we see a person walking, because we can observe movement, activity and change directly, as we are doing with shapes and colors. In fact, we directly observe processes, as walking is not merely movement; it is a specific organized sort of movement, which we call a process or a procedure (Galton, 2006a). For more clarification; when we see a person walking at one instant of time (in a snapshot of the SNAP ontology); although we can see the person fully, the walking is missed in the snapshot, because the person is motionless and wholly present in the snapshot. At best based on the posture at the moment, we can say that person’s posture has characteristics of sequence of postures of a walking person. Simply put, we might presume (with probability), that a person is walking, yet the walking itself cannot observed since it is not in fact existing in the snapshot picture. We may argue that if walking cannot be encompassed in a SNAP ontology, a world that has walking cannot entirely consist of SNAP ontologies, therefore the SPAN ontology is needed to place the walking in it (Galton, 2006a).

Events are perhaps the most extensive information container for dynamic geo-historical phenomena. In order to explain any event well enough, we should take into account its objective and results, its individual participants, its place in space and time, and its relationships to various other events. Indeed representing enough large number of events along all these dimensions may enable us to analyze and discover underlying social historical processes of the globe (Grossner, 2010).

3.3.1 Event and Object

We are typically considering the globe as some entirety of things. Things or objects are separate well-defined pieces of reality. Explaining the world with objects and things is very normal for us, since our language and our technical tools are extremely object-oriented (Galton, 2005). In the real world, objects are defined as conceptual spatial entities with a variety of characteristics and a set of relations with other things, every one of which may alter separately and in different ways (Galton, 2005). Although objects have often been used in GIS modeling, splitting of the concepts has been typically insufficiently emphasized; this has resulted e.g. in mixing conceptual behavior and physical representation of entities (Labbe, 2002).

For instance, if the term 'present' is used in a way that a physical thing can exist both in space and time (Wahlberg, 2009). We can claim that the work desk before us is present “at any time” at which it exists and “at any position” within its spatial location at that moment (Fine, 2008). The commonsense view of objects' persistence is in accordance with the technical metaphysical view referred to as endurantism in philosophical discussions. Physical objects according to endurantism are (Wahlberg, 2009): a) three dimensional, b) persist through time by being "wholly present" at all times as numerically the same entity. "Numerically the same" shares the concept that we are handling with a strictly single identity in time, which means only one and one object exists over time, not two or more objects in series. And the Objects are finally c) changing essentially over time.

The majority of ontologies sharply differentiate objects and events through some key aspects as following (Galton and Mizoguchi, 2009):

1. Relation to time: “an object is present as a whole at each moment of its existence; an event only exists as a whole across the interval over which it occurs”.
2. Nature of parts: “an object can have spatial parts, but does not have temporal parts; an event has temporal parts and may or may not have spatial parts”.
3. Change: “an object can have different properties at different times, and is therefore able to undergo change; it does not make sense to speak of an event changing”.

These three differences are highly interrelated, which means it is difficult to accept any one of them without consequently accepting the others. They together, define the differences between continuants and occurrents which are widespread in both philosophical as well as ontological works. The terms endurant and perdurant have been sometimes used synonymously with continuants and occurrents. Objects and events show some essential interrelated similarities despite these substantial differences (Galton and Mizoguchi, 2009):

1. Discreteness: “objects and events are discrete individuals which may be referred to using count nouns”. When it comes to events, they may be nominalisations of verbs, such as a run or a walk, however oftentimes they are not, such as a fight, an accident or a crash.
2. Non-dissectivity: “the parts of an object or event are not themselves objects or events of the same type, e.g., half of a chair is not a chair, the first half of a walk to the station is not a walk to the station”.
3. Definite extension: “objects and events have well-defined extensions: an object occupies a region in space, an event takes up an interval in time “.

It is important to note that, the properties of objects have to be interpreted with regard to space whereas properties of events should be interpreted with respect to time. By doing so, we can distinguish objects on one hand as 'spatial' entities with events on the other hand as 'temporal' entities (Galton and Mizoguchi, 2009).

3.3.2 Event and state

The notion of causal relationships for events and states goes many years back. Mill, (1843) discusses the instance of a man passing away due to consuming a specific meal. Keeping in mind that eating such a meal is not inevitably deadly, he declares that the consuming of the dish could not be the reason for the death, yet instead merely one among a number of reasons – for instance connecting to the man's wellbeing state – which together triggered the fatality. Particularly, he keeps in mind that “the various conditions, except the single one of eating the food, were not events but states”, and uses this to clarify why we often consider the event as the cause as opposed to the combination of the event and other states. As described in Davidson, (1967), as a matter of fact, the consuming of the meal is the entire cause, however its effect relies on the existence of 'standing conditions' which are not indeed causes. Mill considers conditions as components of the cause, but Davidson differentiates the event as the cause from the state as the preconditions which enables the cause to have its effect (Galton, 2012). Steward, (1997), similarly, is troubled by the usage of states in the 'causal networks' who has suggested for modelling causal mental relationships (Galton, 2012). The idea of states being causally effective undoubtedly raises some problems. States are in general passive, whereas for creating an effect things need to be active. This indicates that the idea of a causally efficacious state is problematic and should be further analyzed (Galton, 2012). Davidson insists that the causal relationship between events need to be independent of how we explain them (Davidson, 1967). “If states are too tied to language, they would seem to partake of the nature of facts” (Galton, 2012). Steward, (1997) says “it might be said that since facts are not part of the natural world at all, they cannot . . . be said to bear causal relations to anything at all”. To understand the role of states in causality, in this thesis we use the notion of state in which change from one state to another is considered as an event. So, the life of an object can be represented through series of state-event sequences. This notion allows us to say, for instance, that event E simultaneously initiates state of S_1 and terminates state of S_2. This notion appears to be fairly natural, and at the same time indicates an ontological commitment to such points as being in a state of S_1. It can have at least two type of relationships to events; ‘initiated by’ and ‘allows’. For instance; state S_i is initiated by event E_i or state S_n allows event E_n to happen (Galton, 2012).

3.3.3 Event and Process

We typically consider the world as a totality of things, likewise we often think of the world's history as a totality of events. However, similar to objects, most of the things that take place in the world are not packaged into distinct events (Galton, 2005). There is a continuous flux, the wind and the rain, the consistent motion of people and animals, the growth and degeneration of plants, the flow of streams and sea currents, and the concomitant erosion and deposition of soil, every one of which are spread out through time in constant variation. Hence, only with mindful artifice we can differentiate some pieces of sufficiently distinctive characters to the description of events. For such temporal activities, the notion of ‘process’ has often been used as a broad term, but “it is a notoriously slippery term whose meaning has proved hard to pin down” (Galton, 2005, 2000).

Different authors have different definitions for the process. Vendler, (1967) proposes “Running, writing, and the like are processes going on in time, that is, roughly, …they consist of successive phases following one another in time. … Running and its kind go on in time in a homogeneous way; any part of the process is of the same nature as the whole”. Pustejovsky, (1991) defines a process as “a sequence of events identifying the same semantic expression”. He says “The verb walk as used in ‘Mary walked’ denotes an activity of indefinite length. That is, the sentence itself does not convey information regarding the temporal extent of the activity”. Moens and Steedman, (1988) conclude that sentences such as ‘Harry climbed’ “typify a third aspectual category, which we will call for obvious reasons a process. Most utterances of such sentences describe an event as extended in time but not characterized by any particular conclusion or culmination”. Allen, (1984) says “Processes refer to some activity not involving a culmination or anticipated result, such as the process denoted by the sentence ‘I am running’”. Sowa, (1999) says: “A process is an evolving sequence of states and events, in which one of the states or events is marked current at a context-dependent time called ‘now’”.

Some authors have defined the difference between a process and an event in duration. Thus Salmon, (1984) says: “The main difference between events and processes is that events are relatively localized in space and time, while processes have much greater temporal duration”. “In space-time diagrams, events are represented by points, while processes are represented by lines” (Salmon, 1984). Sylvan, (1992) states “Events themselves are not normally regarded as processes, as they do not go on, but happen and are finished. But no doubt under a different stretching of the term process, events may be encompassed, as point or short duration processes.” Galton and Mizoguchi, (2009) state “Boundedness is a precondition for the assignment of any definite duration: processes endure, but only once we have assigned bounds to them can we speak of duration, and the act of assigning bounds means that we have switched our attention from the process to an event”.

The differences between processes and events can be summed up under two headlines (Galton, 2006a); first “how processes differ from events”, and second “how they are related to events”. Processes differ from events; as they can be experienced directly, can undergo change, are open-ended and are homogeneous, whereas events lack these properties. Indeed when it comes to properties such as the ability to be directly experienced and the ability to undergo change over time, processes are more similar to objects than to events. The processes all around us are changing, e.g. the running can get faster, or become slower. Consequently traditional differences between continuants and occurrents become more distinct between the world of direct experience, made up of objects and processes, and the world of historical records, made up of events (Galton, 2006a). “A corollary of this is that our snapshot of the world at one time must contain processes as well as objects; snapshots are thus not static but have an intrinsic dynamism which may be thought of as providing the ‘power source’ for the generation of events” (Galton, 2006b). This is in contradiction with the view of Moens and Steedman, (1988) that consider a process as a kind of event.

Still, events and processes are closely related and they can be explained with regards to relationships they have to each other (Galton, 2006a):

An event can be defined with regard to:

- its constituent processes; e.g. "He had a swim", and
- initiated or terminated processes by it; e.g. "He began/stopped swimming".

A process can be described:

- with regards to an event, which it is a constituent of; e.g. "He is swimming a length".
- as the open-ended repeating of an event; e.g. "He is swimming lengths".

Through these relationships the algebra of processes and events can be extracted; where “constituency, initiation, termination, and repetition figure as operators for converting process terms to event terms and vice versa” (Galton, 2006a).

3.3.4 Causal Relations amongst Events, States and Processes

Considering the notion of causality in modeling, the evolution of dynamic systems is of vital importance. For instance, “an accident on the road causes an obstruction which causes reduced traffic flow. The accident is an event, the obstruction is an object or a state, and the reduced traffic flow is a state or a process” (Galton and Worboys, 2005). However, “causes” are not the only causal relationships that are of interest for modelling. Worboys and Hornsby, (2004) list different causal relations for event-event relationships such as “initiation, perpetuation, hindrance, and termination”. For event-object relationships, they list “creation, sustaining in being, reinforcement, degradation, destruction, splitting and merging”. Galton and Worboys, (2005) examine some of these relationships more closely from the viewpoint of events, states, and processes to derive a more systematic classification of the relationships. They advocate: “states enable states” (It would be false to say that states cause states). States perpetuate or maintain states. States are terminated by events. A state is terminated by an event which means originating of a state incompatible with the previous state. Events cause events, which in turn can cause other events. The perpetuation relationship exists between processes; however we can have a similar relation between a process and a state. Events initiate states.

Indeed, each of causal relationships is specific to certain combinations of kinds, as following:

- Event-Event: causes
- Event-State: initiates, terminates
- State-State: enables, perpetuates, disables
- State-Event: allows, prevents
- Event-Process: initiates, terminates
- Process-State: maintain, enables, disables
- Process-Process: perpetuates

3.4 A conceptual event-oriented model proposed in the study

This sub-section is dedicated to the proposed generic and novel event-oriented model for mapping dynamic phenomena of the world. As discussed in the previous sub-sections, the essential concepts needed for modeling dynamic phenomena are including objects, states, processes and events. These three main concepts are the basic components of the proposed event-based model. Figure 3-2 illustrates the schematic concept of the developed event-centric model. In this model, the term “Property” has a significant role. As “Property” we refer to semantic functions, spatiotemporal characteristics and causal relationships between events, states and processes.

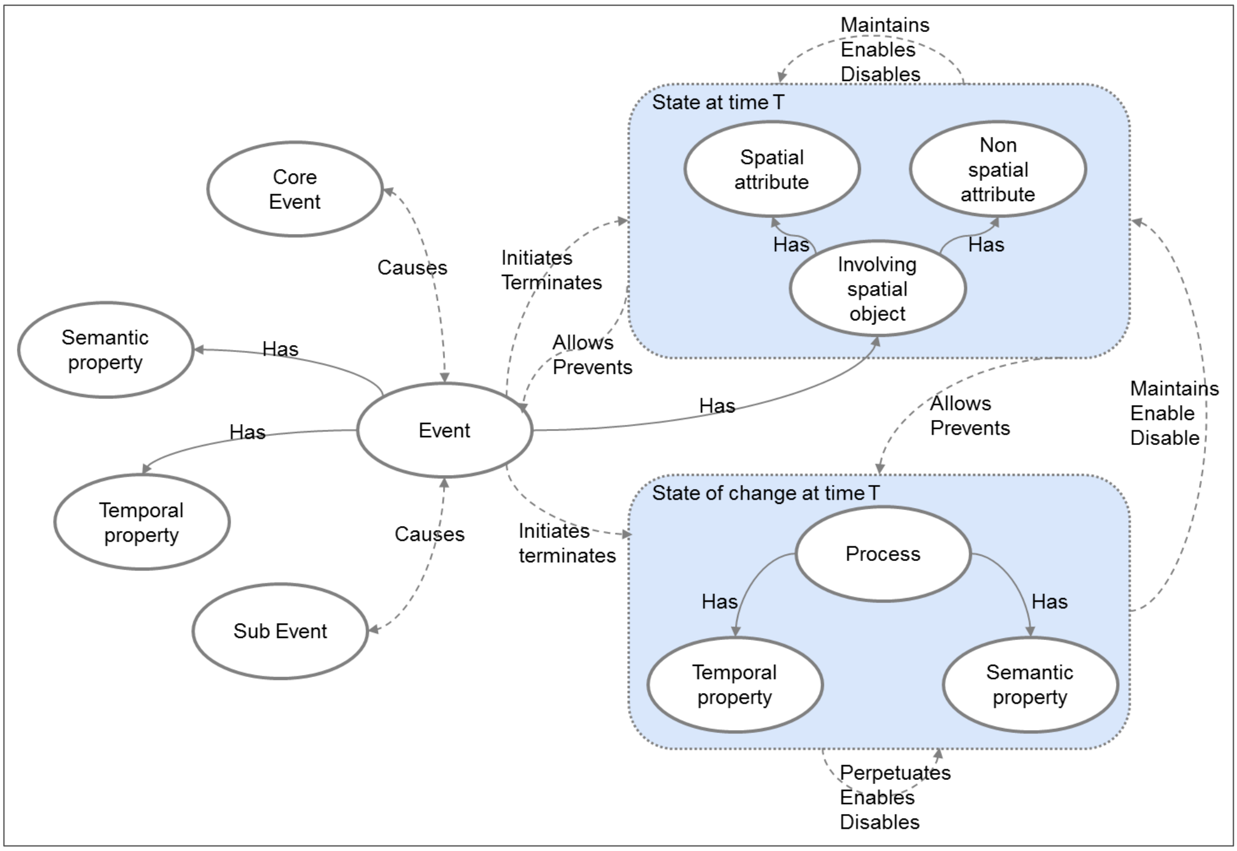

Figure 3-2: Schematic concept of the proposed event-centric model.

The general picture is as following; objects belong to states, and states and processes belong to the 'dynamic snapshot' view of the world at one time. As our snapshots contain both processes and objects, they are not anymore static but have an inherent dynamism which provides the foundation for generating events. By looking at the snapshots we can see different objects, in various states, which undergoing particular processes. These snapshots are constantly renewed as time passes; the snapshots alter from one moment to the next, because the present elements in the snapshot can undergo change. Whereas, events are fixed historical records, which are not renewed or replaced by a new record, however as time passes and events happen, they are gradually added to the record. The model is providing a standard way to mathematically model the changing world and a firm basis for the logical modelling of dynamical systems due to considering processes in the snapshots.

The proposed framework is scale independent and can be used at different granularity levels. Indeed, a process at a specific granularity level may be a sequence of events at a finer granularity. However, sub events of an event should be exhaustive, i.e. all processes of any event should be represented in its sub events. Hence, at any specific granularity level that does not have any sub event only processes of events should be able to represent the full picture of processes happening. In other words, if we consider network of events as a tree, only processes that are in leaves should exhaustively represent all happening processes at the finest granularity level.

Furthermore the model can be used for mapping very different types of events, geographic events, moving objects or events such as football matches or events in the human body. The next few examples can simply illustrate how the framework can be used at different scales for different types of events, and what are different elements of the model and their causal relationships.

Example 1: Analogy between a human body and a city

The analogy between human body and cities has been an inspiration to develop the conceptual event-oriented model in this thesis. Despite obvious differences between cities and human bodies, there are many interesting similarities; both are living and evolving organisms that grow or decay and external or internal factors can affect them. There is a material and energy flow system in cities similar to human veins; free flow is desired and blockages are undesirable in both systems. Both have most of the time local and centralized management systems – if considering the conscious mind as central management and most of sub conscious processes such as healing a wound as local management – and both have sensors and agents throughout their entire system. The anatomy of human body is indeed like a city map; it has cells or organs as its objects with spatial and semantic information.

However, to prescribe for a human body or make decisions for a city, aside from a body anatomy or a city map, clear understanding of events happening in each system, states of the systems, and effect of different interventions in the systems are required. These effects cannot be fully analyzed or predicted, if happened events and involving processes from one state to the next state are not understood well enough.

By considering a heartbeat as the event of interest our model will have following components:

- Event: heartbeat
- Core event: electrical impulses …
- Sub-events: contractions of the myocardium, relaxations of the myocardium …
- Processes: muscle contractions, muscle relaxations, flow of blood …
- Involving objects: heart, blood …
- State: the state of all involving objects such as heart muscles and blood at each moment

For each process like contraction and relaxation in the heartbeat event, same sub-events should be presented to analyze the event at a finer granular level. However at any granularity level, one can ignore all sub-events and just consider processes to analyze the event of interest or other core events. It is clear that by understanding relationships between events and core/sub-events as well as processes that happen at each level, we can easily use the model as a mathematical formula to predict future events under specific circumstances. However, to reach this goal, relationships between states and processes should be known. For instance if we know that a muscle with a blockage in its veins (state) may distract the contraction of the muscle (process), we can predict that in case of a blockage in a myocardium, heart problem may happen. These predictions can go even much further; for example, by analyzing core events that can cause a blockage of a muscle or analyzing core events of those core events and so on, we can go back to the root of a problem and try to prevent it at its very early stages. This shows the prospect and potentials of having a comprehensive event model for a city or any geographic system; we would have physicians for our cities that can help city planners how to plan changes in future or how to heal a diseased city. A main pre-requisite to reach this goal is to develop a solid ontology for events and involving states and processes.

Example 2: A football game

We consider a football game as our second event of interest. Following components are present in almost any football game:

- Event: a football game
- Core event: champion league
- Sub-events: a player scores a goal, referee whistles, audience cheer …
- Process: whistling, running, kicking, shouting, slipping, falling …
- Involving objects: players, audience, referees, stadium …
- State: the state of all involving objects at each moment

All these components may have some attributes; spatial, temporal or semantic. For instance the event itself has a temporal property; its date and time, and a semantic property such as its category as sport. The attributes of states (including objects) and processes may have both variable (time-dependent) and constant (time-independent) properties, whereas event attributes are all constant and time-independent. E.g. the spatial attribute of players change in each snapshot but their name as a non-spatial attribute remains the same. Analyzing the causality between different concepts is described below:

- Event-Event: a player tackles which 'causes' another player to fall
- Event-State: an unfair act of a player 'initiates' a state of anxiety in other players
- State-State: wet turf 'enables' grass to stay green
- State-Event: wet turf 'prevents' players from scoring a goal
- Event-Process: After first goal audience 'started' screaming
- Process-State: Players run out of the playground to 'stay' warm
- State-Process: wet turf 'allows' players to slip
- Process-Process: so long as players keep running around the football field after the victory, people keep screaming further.

Example 3: Octoberfest

We consider Oktoberfest as our third event of interest. Following components are present:

- Event: Oktoberfest 2015 in Munich
- Core event: King Ludwig I and Therese of Saxe-Hildburghausen wedding on October 1810, municipality official decision to hold 2015 Octoberfest event …
- Sub-events: constructing arena, hold fairs, deconstructing arena …
- Process: construction, holding the fair, disassembling tents and arena …
- Involving objects: local breweries, tents, construction labor force, organizers, participants in the fair …
- State: the state of all involving objects at each moment

This event has three main sub-events, each of which can have multiple sub-events with different timing. These events can be theoretically broken down to any granularity level. For instance; each sub-event may have following sub-events as well;

- 'Constructing arena' may include; constructing infrastructure, constructing tent A, constructing tent B …

- ‘Constructing tent A’ may include: make contract with the constructor, transport required tents and material, hire people to build the tent, install tents …
- ‘Install tents’ may include: dig a hole, raise the main column …
- ‘Raise the main column’ may include: connect different pieces of rods to build the column, raise the column with crane, fix the base to the earth …
- ’Connect different pieces of rods to build the column’ may include: screw rod A to rod B …. The process for sub-event ‘screw rod A to rod B’ is screwing.

The approach can be used to model events at any scale; the core or any of the sub events of examples can be for instance the event of interest. Defining the granularity level of the model depends on the purpose for which the model is used and availability of data. A municipality official may need to know only start and end date of each of three sub-events of Oktoberfest, but a blue collar labor should know in detail how much time he or she requires to connect to rods.

4 AN EVENT-CENTRIC FRAMEWORK FOR CAPTURING AND STORING EVENT INFORMATION

In this chapter we present a framework that provides the possibility of collecting and storing event-related information in a geo-referenced format on OSM platform. Four mechanisms were established to collect event data; among them two mechanisms facilitate volunteers to manually gather information and two mechanisms collect and store data automatically from the web. Due to the central role of the event notion, the introduced event definition in the previous chapter is based for the technological development of the work. The OpenStreetMap (OSM) platform, as one of the most prominent examples of spatial user generated data, was selected to store the event information because the objects inherently are geo-referenced (main problem with user generated data) and the semantic information can be assigned to them through tagging. This gives OSM platform a unique potential to collect and store the geographic and semantic information.

4.1 OpenStreetMap

"The OpenStreetMap project is a knowledge collective that provides user-generated street maps", which was founded by Steve Coast at the University College London (UCL) in 2004 (Haklay and Weber, 2008). The removal of Selective Availability (SA) of the Global Positioning System (GPS) by President Bill Clinton in 2000 (Clinton, 2000), as well as technological advances such as developing affordable GPS devices and publication of the interchange standard (GPS eXchange/GPX), were indeed prerequisites of the OSM project. Furthermore, advances in positioning, web mapping, wiki technologies, and increased bandwidth enabled experts and amateurs to easily create and share geographical information (Coleman et al., 2009; Goodchild, 2007b, 2007c).

4.1.1 OpenStreetMap community

OSM follows the production model of Wikipedia; which aims to create a set of free to use and editable map data licensed under new open copyright schemes. There are three categories of OSM contributors; a significant number of contributors create and modify the map of the globe collaboratively using the OSM technical infrastructure, a core group of volunteers dedicate their time to developing and enhancing OSM's infrastructure, including server maintenance, writing the core software program that handles the transactions with the server, as well as creating cartographical results. The third group of contributors is a community of software application developers who create software programs to make OSM information available for further usage throughout various application domains, software platforms, and hardware tools (Haklay and Weber, 2008).

The common procedure to create a new feature in OSM is first to gather GPS tracks in the field, upload them on the OSM server and ultimately modify them with either a web-based service easily accessible from the OSM user interface or with a more advanced standalone software application like Java OSM (JOSM) editor. Another approach to input data is the digitization of desired features based upon an aerial imagery. Nevertheless, this can be done just with certified sources of data compatible with OSM such as from Yahoo! Imagery, or from many governmental mapping agencies such as the United States Census, or some commercial firms such as AND, that have given OSM access to their base map datasets. After creating the spatial data, different attributes are added to it to describe the features. In theory, it is totally free to add attributes, but the majority of contributors adhere to an attribute structure specified collaboratively by the OSM community.

As of July 2015, OSM has more than two million registered users which increase steadily, and a data contribution rate that continues to rise very quickly (OSM Community, 2015). The OSM dynamic community organizes different social events and map parties and a yearly OSM conference such as State Of The Map. This has made OSM a symbolic VGI project, which illustrates the power of VGI and attracts more and more new users. For the well mapped locations of the globe, it provides an alternative to commercial datasets whereas for areas of the world that are not well mapped due to the fact that they are not considered profitable by commercial mapping firms, it in some cases provides the only available mapping (Haklay and Weber, 2008).

4.1.2 OpenStreetMap elements and database structure

OpenStreetMap (OSM) is a collaborative volunteer based mapping platform. The large amount of data is stored in a robust key (tags) value pair database. The database can store any sets of key values in the form of strings. These functionalities have made OSM a powerful community for collecting both geometry and semantic information of physical objects. The main OSM database implemented in MySQL lies at the heart of OSM's technological infrastructure, which holds the real-time information. The database schema is developed to support various wiki actions, such as versioning and rollbacks, and keeps duplicates of changed or deleted features and attributes indefinitely (Haklay and Weber, 2008).

OSM uses a simple data structure with four core elements; nodes, ways, relations and tags, which are stored in database tables on a PostgreSQL Server. Each row of the database contains key-value pairs, so any point in the database can be related to several attributes. A node is a solitary geospatial point characterized with a unique latitude and longitude coordinate. A way is an ordered listing of 2 to 2,000 nodes which either represents a continuous path on the map or a closed way or polygon, if the very first and the last node are the same. A relation is a logical grouping and collection of several elements such as nodes, ways and relations standing for an advanced concept such as a rout or a turn restriction in a road. The definition of the relation is specified by the tags which provide semantic information and describe the elements to which they are attached by different properties (OSM Community, 2015). Indeed all non-spatial information including names, classification types, and other variety of properties are stored in the form of tags.

OpenStreetMap database is very simple; a piece of data – a row in the database – consists of a pair of (key, value). Using this concept, for any point in the database, multiple metadata can be associated as the point's attributes. This system allows all different kinds of data to be entered and accommodates different terminologies used in different parts of the world. Figure 4-1 shows an exemplary key, value data for a node in OpenStreetMap using JOSM offline editor.

Figure 4-1: Example of key, value data for a node in OpenStreetMap.

4.2 System Architecture

The general convention in OSM community is to store spatial data. There have been discussions regarding storage of temporal information but it has been discouraged. Polous et al., (2013) also revealed that there is a lack of event detection and pattern recognition possibility for OSM community. Therefore, a continuous maintenance is needed to update the map for each period of time for involved objects that events are happening on them. Mapping temporary objects of repetitive periodical events affects the quality of OSM for the users who are not frequently updating their downloaded maps (Polous et al., 2013). However, it is also almost impossible to update the map for irregular events with very short cycles of repetition.

In this study, the OSM database and technology infrastructure is extended to explore the idea of temporal data collection and storage in the same OSM format. Events are collected and entered into the database using special tags. OpenStreetMap is powered by a combination of Open Source software that can be downloaded freely from the internet. Figure 4-2 shows different components of OSM architecture in green and few other open source APIs. The components in blue are additions to the OSM framework to handle temporal information of events in this thesis. The framework has three main segments; a) the collection mechanisms for the gathering of temporal event information (shown at the top of the Figure 4-2), b) the server infrastructure (shown at the middle of the Figure 4-2), and c) utilization mechanisms and volunteered Location Based Services development for the collected event information (shown at the bottom of the Figure 4-2). The first two parts of the framework will be described in this chapter and the final part will be explored in the next chapter.

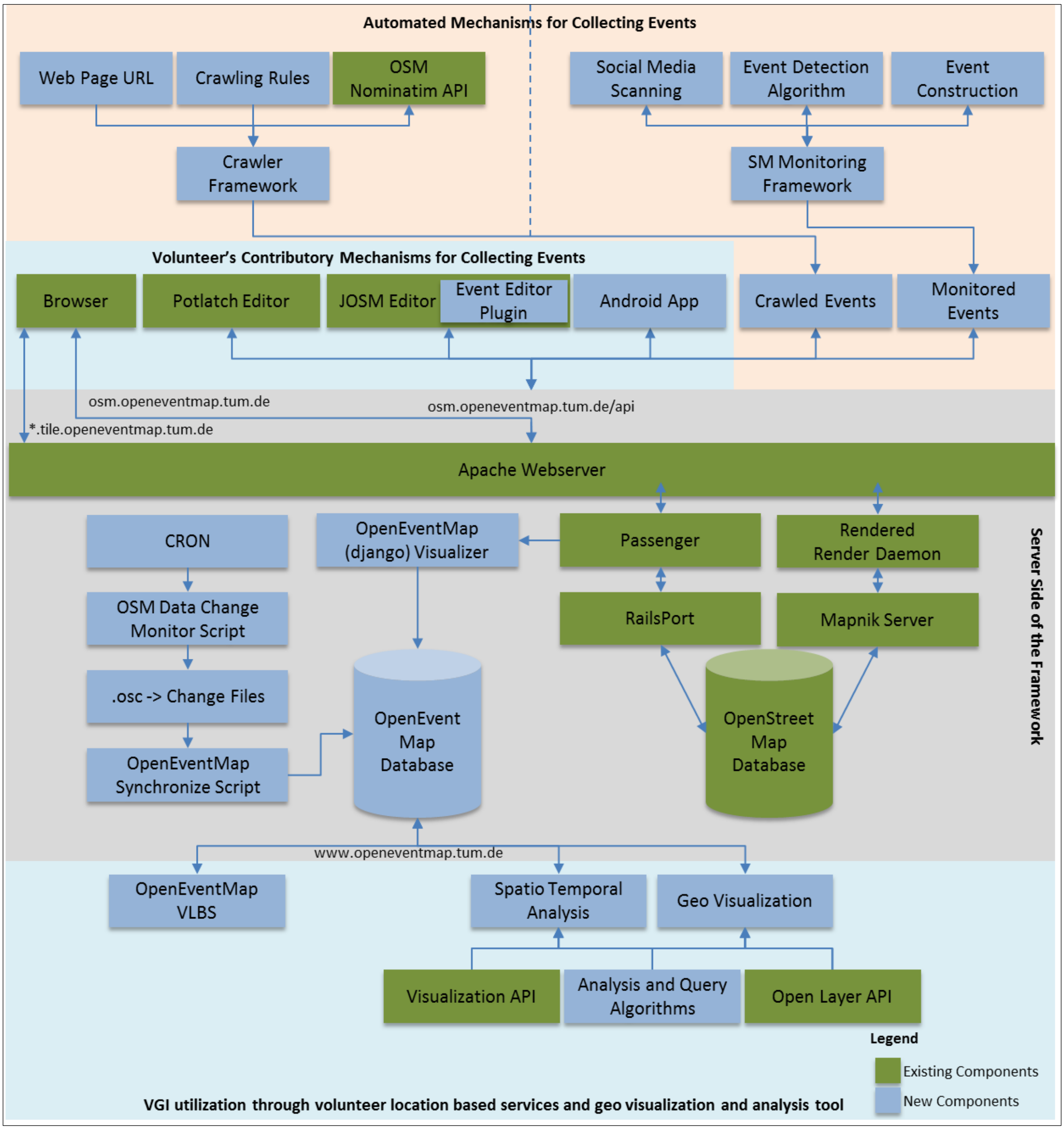

Figure 4-2: Extended OSM framework for managing temporal information.

4.2.1 Saving temporal information in OSM

Figure 4-3 shows four mechanisms that have been developed within this thesis to facilitate the collection of events data either by contributors or automatically; a) a plugin for JOSM offline editor, b) an Android application, c) a web crawler, and d) a social media monitoring framework. The collected information is stored in a local database (so called Event-OSM database). The database is compatible with the original OSM database and is frequently synchronized with OSM server. The details of event based tag metadata, system architecture, and technical details of the entire system are presented in this section.

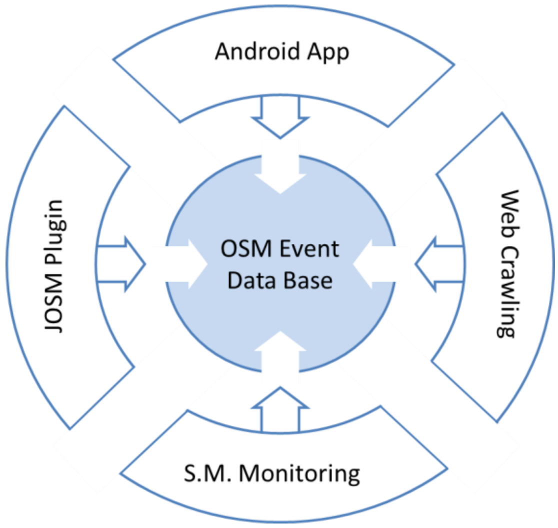

Figure 4-3: Four data entry mechanisms to the OSM event database.

4.2.2 Describing Event Attributes through Tags

In the context of this study, the event refers to any change that happens in spatial and semantic information of object(s), which are bound to a location at a certain point or specific period of time. Here, events are modeled with five attributes:

1. "What" for all desired semantic information like the type or class of events
2. "Where" for the location of events by assigning objects that are involved,
3. "When" for the time point or duration of events,
4. "How" for the pattern of occurrence
5. "Who" for the organization, agent, or external force.

OSM has predefined set of tags to describe spatial data. The challenge was to come up with a set of tags that can describe temporal attributes of given points adequately. This set of tags should be easily fit into the existing system, not just for describing attributes but also for other purposes like searching event elements. We associated events primarily with one of the nodes, ways or relations. This is done by associating special event tags to the OSM elements. The developed syntax for event related information in OSM platform is as following;

- event:<event_index>:<event_attribute> = value
- event_index belongs to [0,1,2,3,..,N]
- event_attribute is one of the values like name or event category

The primary tag to mark any element as event is to set the following tag;

event: "YES".

In total, we consider eleven other tags to associate more information to the OSM elements for event mapping:

1. Name: Name.
2. Category: A list of primary categories user can choose from.

3. Sub category: Each primary category has a list of further sub categories for refined classification.
4. Organization: Main organization related to the event.
5. Start date: Actual day when the event occurs.
6. End date: In case of a single day event, it is the same as start_date but different if it is a longer event.
7. Url: A website that explains more information about the event.
8. Number of participants: An estimate of the number of participants in the event.
9. How often: Event repeat frequency.
10. Comment: Any extra note a user can associate with the event.
11. Related_items: A comma separated list of other OSM elements related with this event.

Figure 4-4 shows these tags in the context of the developed conceptual event model which was presented in the previous chapter. All aforementioned tags cover only the blue colored elements of the model in Figure 4-4. No tag was considered for the gray colored elements of the model, as it could cause confusion for volunteers. These elements can be easily extended based on required information to be collected such as core or sub events and processes. However, for mapping relationships between different elements of the event model a further detailed study is required to refine the syntax.

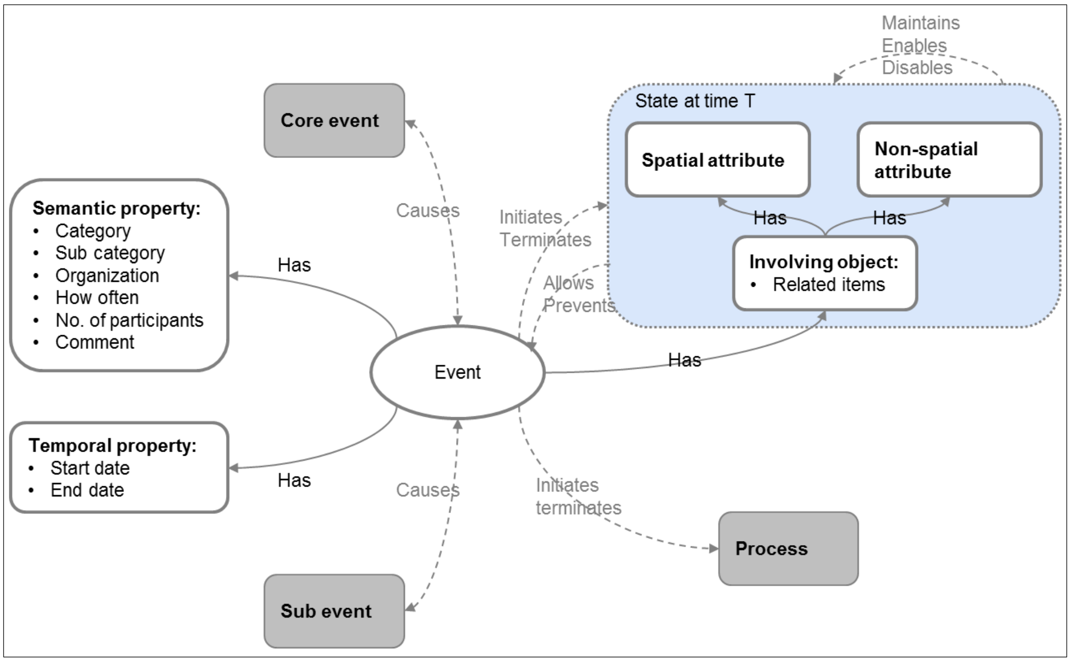

Figure 4-4: Suggested tags in the context of the developed conceptual event model.

Figure 4-5 shows an exemplary event which has been mapped using the developed syntax for OSM.

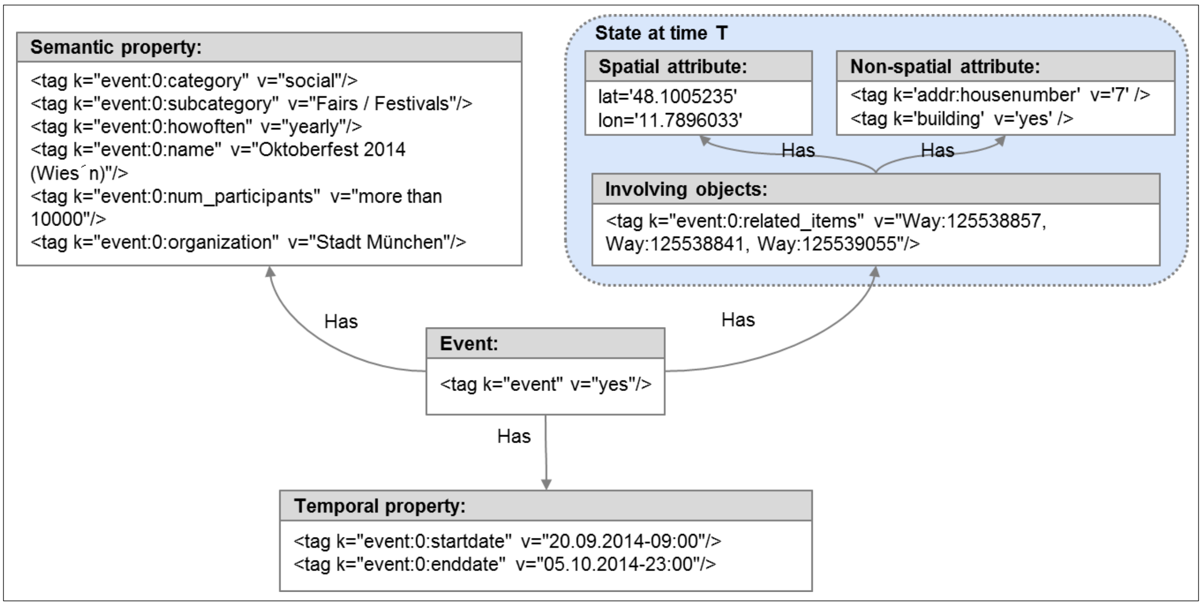

Figure 4-5: An exemplary event mapped with the developed syntax in OSM.

To fully facilitate the mapping of the N to M relationship between objects and events, two extra strategies were considered in developing the syntax. The first strategy is; namespacing in tags and assigning an index to each event. This enables contributors to assign any different number of events to the same location or object, which is very common for locations like exhibitions or sport arenas. The character ":" is used for namespacing. The format of any event tag is as follows:

event [namespacer] event_index [namespacer] event_attribute = value,

where event_index starts with 0 and increases by one for each event, event_attribute can be one of the defined tags in the previous sub-section. Figure 4-6 as an example shows a way (a pedestrian street) on which, two social events at two different periods of time are happening.

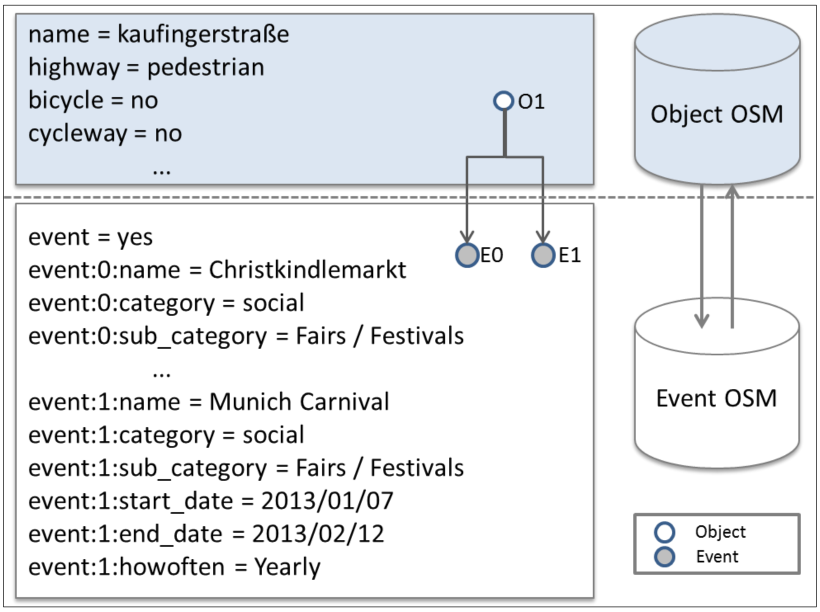

Figure 4-6: Demonstration of N to M relationship between object database and event database; one object and two events.

In contrary with the aforementioned situation, sometimes we have to map an event, in which many separated spatial objects are involved, like brewery tents for Oktoberfest. In this case instead of creating different events with the same semantic information for each tent, we create just one event node and add all tents (OSM objects) to that event as its involving objects (Figure 4-7). This will avoid a high redundancy in our data base. The objects in red are selected as involving objects for the events. The developed plugin can automatically retrieve the OSM ids for objects and assign them as related object for an event.

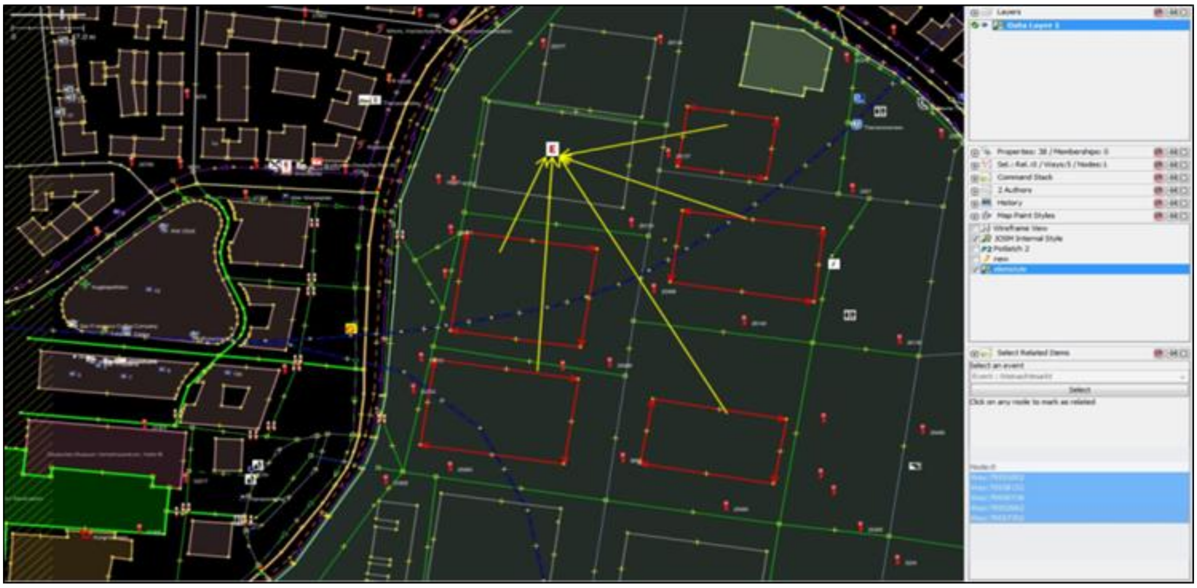

Figure 4-7: Demonstration of N to M relationship between object database and event database; assigning more than one object to a single event.

4.3 Capturing contributed event information through JOSM

To facilitate the data acquisition procedure, an event editor plugin, named "Event Editor" was developed for JOSM offline editor. It provides a Graphical User Interface (GUI) to easily enter the event information to the OSM database. JOSM was selected as it is one the most popular offline editors for the OpenStreetMap community (OSM Community, 2015). This plugin was developed based on the proposed event definition in the previous chapter, to adequately store and manipulate semantic information of events in the same OSM structure format. Figure 4-8 illustrates the GUI of the developed plugin for mapping events and collecting the semantic information for the aforementioned tags.

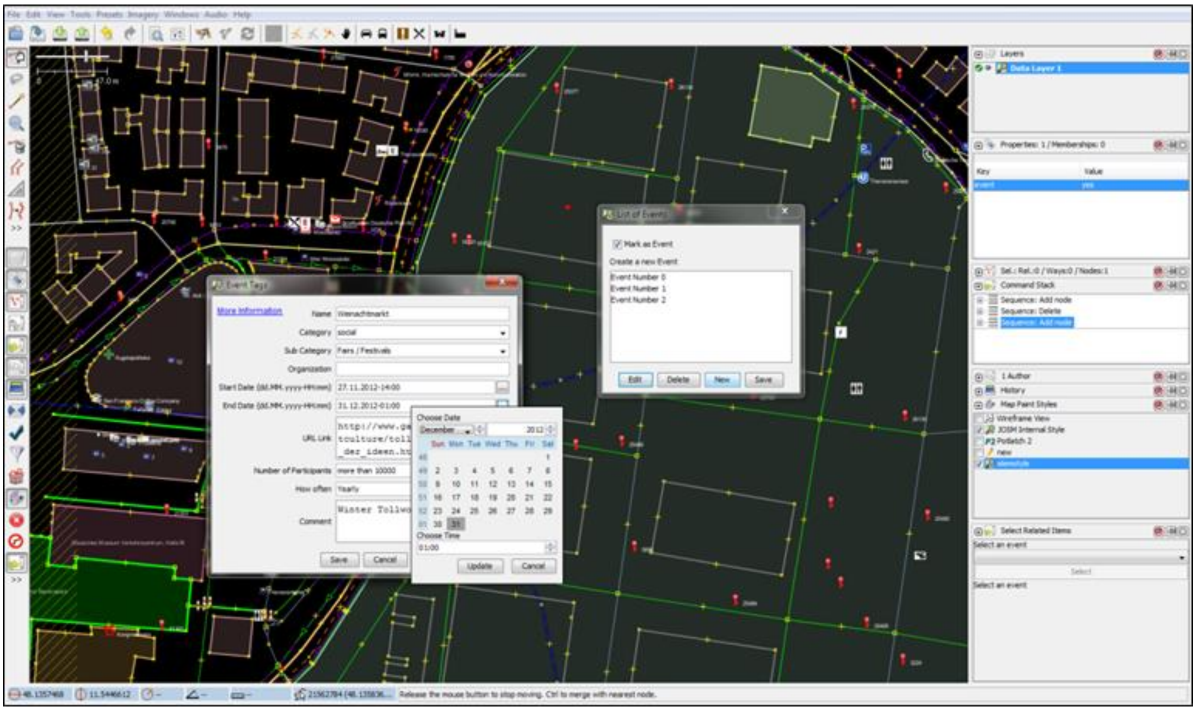

Figure 4-8: EventEditor screenshot in the JOSM environment.

Aside from the aforementioned mechanisms to collect event information by volunteers, there is another way to contribute event information to the database manually. It is also possible to gather the information through the Potlatch web-based map editor which is a free online OSM editor used by many of OSM contributors. It is possible to motivate the OSM potlatch adopted contributors to collect the information through this web-based tool, but the contributors should adhere carefully to the developed event syntax. An example is shown in Figure 4-9. However, developing a special key such as “Event” which has all required pre-defined tags can be a safe mechanism for collecting event information through Potlatch web by volunteers.

Node: 1994284903

Key	Value
event	yes
event:0:category	sport
event:0:enddate	12.10.2014
event:0:howoften	Once
event:0:name	München Marathon
event:0:num_participants	1000 to 10000
event:0:organization	Landeshauptstadt München
event:0:related_items	Way:15805167
event:0:startdate	12.10.2014
event:0:subcategory	Athletics
event:0:url	http://www.muenchenmarathon.de/

Figure 4-9: Adding event information through Potlatch.

4.4 Capturing contributed event information through Android App

In order to ease volunteers and users in adding, editing, mining and visualizing event information on the OSM server, an Android application called 'OpenEventMap' was developed. The application comes with an intuitive user interface that allows users to insert/modify or search/view event information without directly working with the OSM tags. The Android platform can be used in addition to obtain user related information, such as current location.

A wide variety of multi-purpose Android apps have been developed to manipulate the OSM database. For a complete overview of these apps visit the OSM wiki[7]. Vespucci was selected as the starting point for developing the OpenEventMap application, since it provides core functionalities such as navigating a map and uploading changesets to the OSM server. Vespucci interface and its default settings were redesigned and adopted for the purpose of event mapping. For instance some of the Vespucci functionalities were removed from the original application to restrict the usage of the OpenEventMap application for events. The customizations to Vespucci are presented in more detail in the next sub-section.

4.4.1 Android App main activities

Android activities are single. Almost all activities interact with users, so the activities are user interfaces allowing the user to complete one specific task at a time. Although the main activities were already presented in Vespucci, but it is widely extended in OpenEventMap and new activities are designed specifically for the OpenEventMap application.

The main activity user Interface (Main.java) is mostly used to navigate (pan and zoom) the map and display events on the map. Through the action bar the following actions can be performed: a) search events by different attributes such as name, category, start and end date. The search can be performed either in the active bounding box or in a certain radius around the current location (determined by GPS) b) search for events through their places; inserting a query for Nominatim, which allows users to search places by their names c) add new events while enriching them with temporal and semantic information e) upload changes to the OSM server. This functionality is used when new events are created or the existing events are modified and f) undo or delete the created events or changes of any existing event. Figure 4-10 shows the activity workflow of the android application. The main activity is also responsible to utilizing other classes in the background such as:

- Downloading information from the SQL database when the user mine an event search or navigates on the map.
- Downloading information from the API when the user wants to view or edit event information.
- Keeping track of what the user activity
- Initiating the creation of an event
- Providing the possibility of selecting the multiple events for one single element or multiple elements for one single event. In addition, the multiple events at the same spot can be assigned to different OSM elements.

[7] http://wiki.openstreetmap.org/wiki/Android

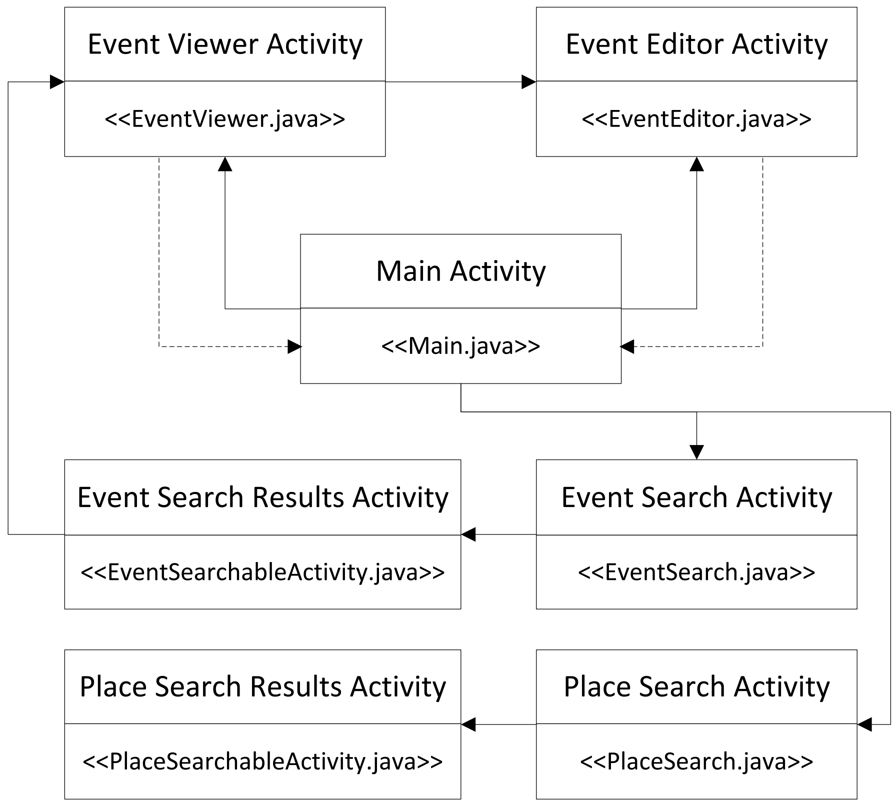

Figure 4-10: Activity work flow: solid arrows show the interaction between current activity and the other activities. Dashed arrows indicate when the Main activity is started.

4.4.2 Android App event viewer activities

The Event Viewer (EventViewer.java) is started when user selects an event on the map. In order to show the most up-to-date information and make changes locally, the event object is rebuilt from the API in the local storage. The UI of the Event Viewer is composed of some Android TextViews displaying all relevant event information in a readable format. Available actions here are:

- Viewing events on map, which brings the Main Activity to show all related items of an event on the map such as in Figure 4-11.
- Starting editing of the events, which starts the Event Editor activity for modifying the current event.
- Adding a new event to the same OSM element, which activate the Event Editor for entering information for the new event.
- Returning to the Main activity.

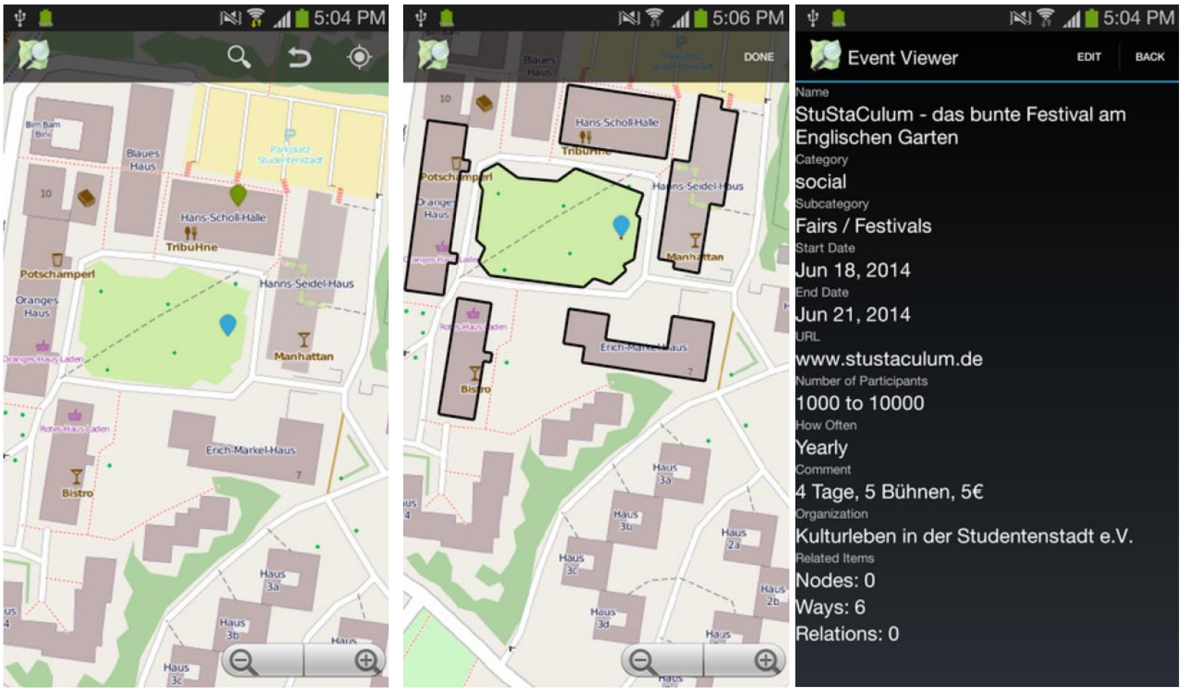

Figure 4-11: Displaying events, related items, and detailed event information.

4.4.3 Android App edit activities

The Event Editor (EventEditor.java) is composed of several Android EditTexts, Buttons, Spinners and Dialogs that allow the user to input event information in a most natural way. For example, dates are entered through a date picker where the day, month and year can be entered in the same intuitive and therefore consistent way as birthdays are entered in the Contacts app. In addition a small calendar could appear on large displays to choose the date directly from the calendar view.

The Event Editor is started whenever the user decides to edit an existing event in the Event Viewer, or when a new event should be created (Figure 4-12). It can be used to perform the following actions:

- Change the values of available event tags.
- Save the changes to the local storage to be uploaded later on.

Start the Main activity to display the map with surrounding OSM elements in order to choose the related items of the event (Figure 4-12).

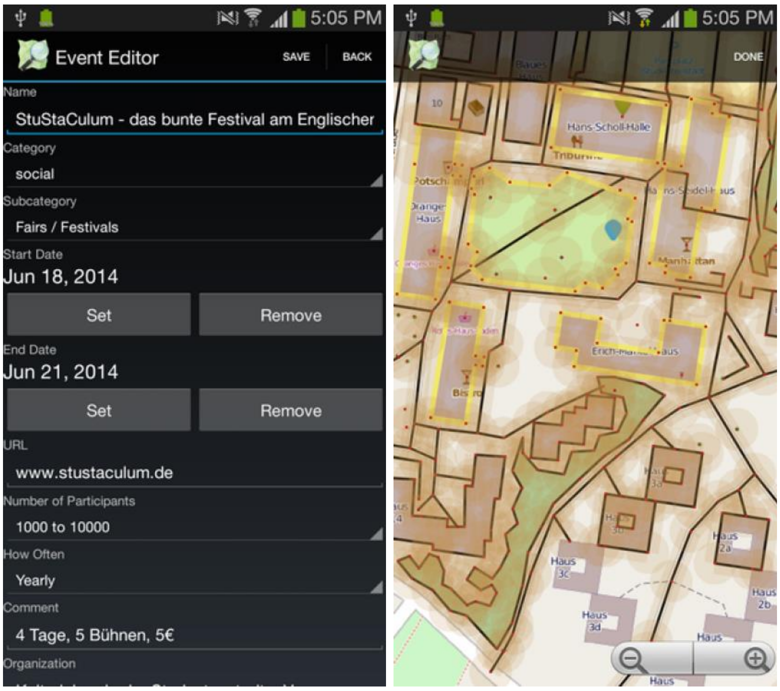

Figure 4-12: Displaying detailed event information, related items of a selected event on the map (ready to be edited by the user).

4.4.4 The Event Search Results Activity

This activity (EventSearch.java) allows the user to search the events based on name, category as well as start and end date of events. Submitting this parameter, another activity is atarted, which performs the search and displays the results in a ListView.

Here the actual search is performed on the SQL database. The results show the most important information about each events satisfying the given search criteria. A custom adapter (EventListAdapter.java) and item layout (res/layout/event_search_list_item.xml) have been created and designed to provide the desired flexibility in displaying the event details. If no matching events could be found the user will have the following options (Figure 4-13):

- Create a new event with already specified information. This saves editing time and makes inserting events much faster after a failed search.
- Return to the Event Search activity and refine the search criteria.
- Stop searching and go back to the map.

In addition, clicking on one of the results starts the Event Viewer for viewing all details about the event.

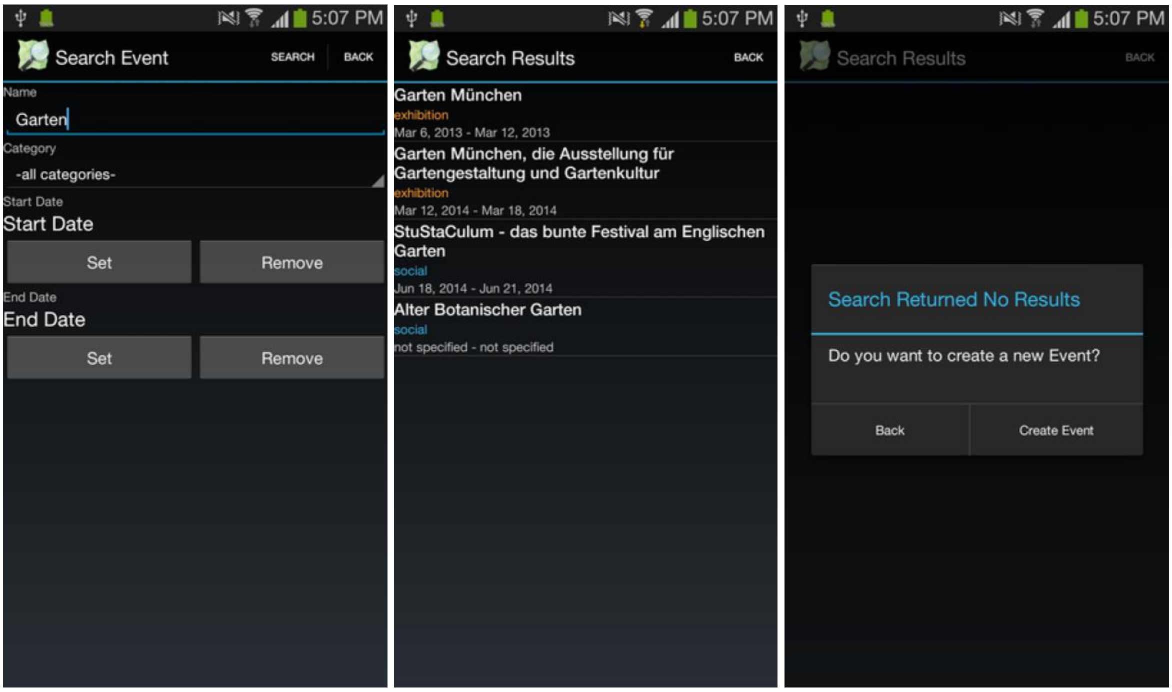

Figure 4-13: Searching events through different criteria, obtaining the information while matching the search criteria, and providing the user possibility of thinking of the next step.

4.4.5 The Place Search Results Activity

This activity (PlaceSearch.java) offers an EditText, which provides the possibility of mining the databases through the Nominatim service. After committing the inquiry, another activity is ran to do the search on the server and display the results. The obtained results are displayed by directly using the result string that the Nominatim is returning. These strings are very long, but also very detailed so that the user can decide which result is actually the one that the user wanted in the first place (Figure 4-14). By choosing one of the results, the Main activity is started again and centers the map on the given place.

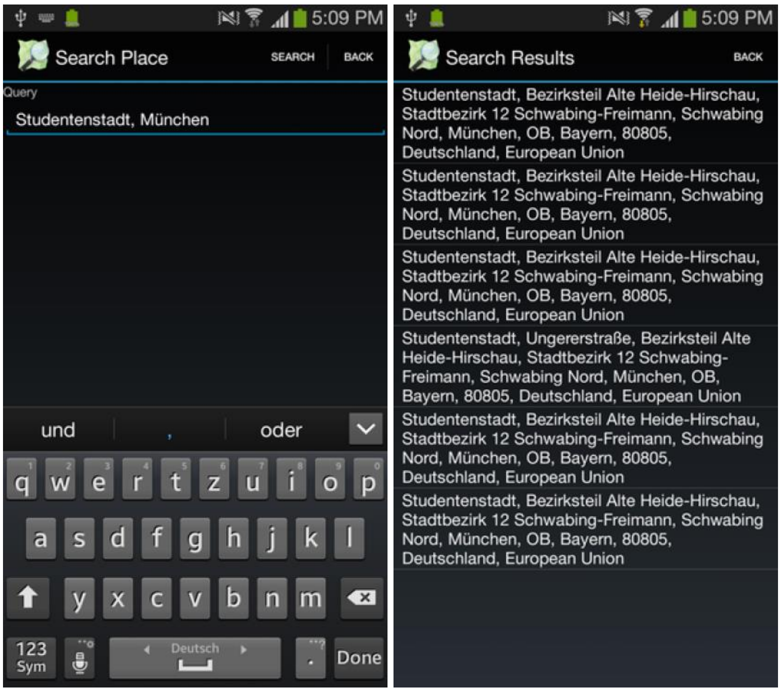

Figure 4-14: Searching for locations through an address or a name, and showing the results for the given criteria.

4.5 Automatic capturing of event information through Web Crawling

For the events, there are viable webpages that announce events information. However the information on these webpages is unstructured and dispersed. In order to abstract event information, and manage them in a more structured way according to certain characteristics such as place, time, event type, a crawling framework was developed to automatically collect this freely available information. Since we only focus on the events information, only the texts relating to the events are extracted from those webpages. Other information, for example, HTML metadata, is omitted.

The crawler includes a generalized core which needs a specific plugin for each webpage. The plugins vary depending on the tags-nomenclature of different websites. Unlike traditional web crawler, this web crawler only focuses on certain webpages and fetching the data according to a given crawling rule. Figure 4-15 illustrates the entire process.

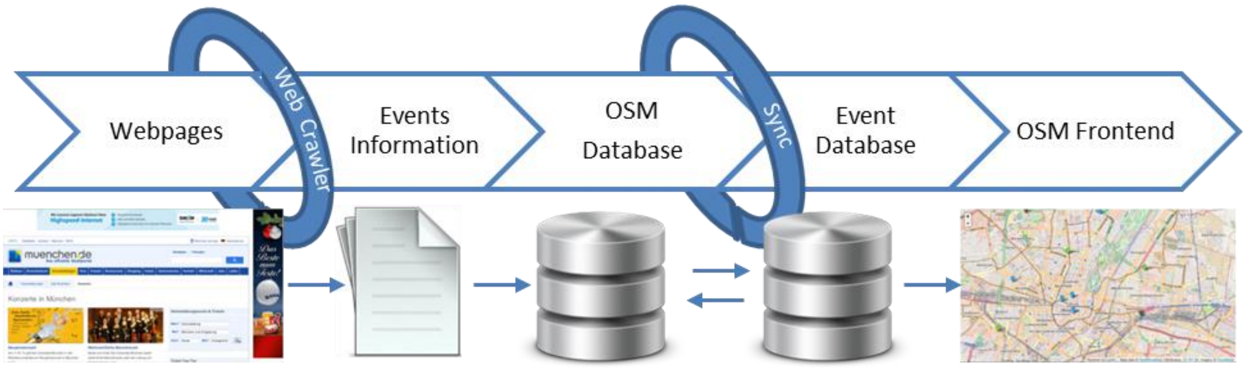

Figure 4-15: Automatic event collection process.

To prevent disturbance on OSM database that is tied with OSM frontend, a local database is used to temporarily store the crawled information from the webpages. The local database is compatible and synchronized with OSM database. Before synchronizing, in order to inhibit redundancy caused by overlapped information existing on different webpages, a data re-arrangement process is conducted in the local database. After data re-arrangement, the synchronization is executed to transfer data from local to OSM event database. Eventually the event information is displayed on OSM Frontend. With a certain set of webpages and their crawling rules, the entire process is fully automated; so that, the events information on OSM Frontend will be always up to date.

4.5.1 Web Crawler

All selected websites use Hyper Text Markup Language (HTML) to present webpages and other contents displayed in a web browser. HTML code is composed of HTML elements. An HTML element usually consists of tags. The tags come in most cases in pairs e.g. and , which are called respectively start and end tag. Between the start tag and end tag, it can be pure text, further tags or other contents (e.g. JavaScript). Start- and end tags usually act as identifiers. Event information, is embedded in the pure text of the webpage's HTML code and is encapsulated by start and end tags. In order to extract event information, it is necessary to find the location of tags, which contain the information of event elements.

First, we select some training sets from HTML content, and manually build crawling rules by analyzing the training set. To build these rules, a so-called 3-which problems needs to be solved; 1) from which tag the information of an individual event begins, 2) which following tags contain the information about the name, location, time, and duration of that event, and 3) in which tag the information of that event ends. Finding the answer to 3-whiches normally faces difficulties. It is mainly because of other available texts in the HTML file such as HTML metadata, web navigation and advertisements. Once the tags and their corresponding text are known, we need to exclude the tags, extracting the pure text and storing them to the right column in the database. This procedure is accomplished in four steps, which is discussed in the following.

4.5.1.1 Downloading the HTML code

Downloading the HTML code is performed by sending a Hypertext Transfer Protocol (HTTP) GET request to the web server. The encapsulated Uniform Resource Identifiers (URI) in the request indicates the web server, to which the request should be sent. As a result, the HTML code of the webpage is retrieved.

4.5.1.2 Manual pattern analysis

By comparing the names of start and end tags in the HTML code, the pure text of tags, and the content displayed in the browser, web structure patterns can be extracted. All patterns extracted here are further referred to as initial rules. Finding the final rules based on the patterns manually is indeed a repetitive process. The initial rules require to be optimized by searching the HTML code for possible different patterns in defining the event tags. It is not very rare to have web pages with different tagging patterns for events. It is necessary to create different final rules for each individual webpage, as different webpages use different tag structures.

4.5.1.3 Fetching information

A java-based framework is developed for the purpose of parsing HTML codes and fetching the useful information. The framework reads the HTML code, extracts information according to the final rules of each webpage and follows a post processing step like data integrity and format unification, if necessary. It includes a generalized core which needs a specific plugin for each webpage. The plugins vary depending on the tags-nomenclature of different websites.

4.5.1.4 Storing the information

Storing the information in the columns of the SQL table of the event OSM database is the last step. Information crawled from webpages that are not listed in the columns of the SQL table is ignored. Each extracted event from a webpage will have one record in the table. If the framework cannot find relevant information for any column in the SQL table, the column value will remain as null for the record. This incompletion can probably be made up by the information crawled and integrated from other webpages.

To start the framework, it is required to feed it with the URL of a webpage and its corresponding final rules. The final rules are composed of patterns that are described by tuples. A tuple is expressed as; [HTML-Tag-Identifier, Tag-Action].

In this way the framework reads the HTML codes tag by tag, once it reaches a tag identifier, it extracts the following pure text in the tag. The second element of the tuple describes the action that should be taken for the tag. Actions are classified into three groups:

1. If the tag is the beginning or ending of certain event information, the action element indicates the start or end of a new record in the database.
2. If the tag contains the event information, the action element indicates the corresponding column name in the local database, where this information should be stored.
3. If the tag is a link that may contain detailed information of a certain event, the action element directs the framework to conduct web crawling iteratively on the embedded link.

Normally before storing the information, data format unification for time and location is required. In our SQL table, the default date format is YYYY-MM-DD. The location of events can also be described in different ways; some webpages use the name of the places while others use a concrete address. To acquire an appropriate unified location format for an event, the OSM Nomination API[8] is used to capture the corresponding OSM identifier of locations. The OSM Nomination API describes the location of events through OSM objects instead of a string address.

4.5.2 Data post-processing

Information crawled from different webpages may sometimes have overlaps or be complementary to each other. In order to reduce the redundancy and at the same time to prevent information loss, further optimization is implemented.

4.5.2.1 Data integration

Records referring to the same event may repeat several times in the database. These redundant records maybe exactly duplicated, or having some common values in some attributes such as name,

[8] http://nominatim.openstreetmap.org

start-date, and end-date or related items. We regard these records in the local database as redundant records. An easy way to deal with these records is to delete them. However, in order to prevent information loss, we integrate these records as one record in the same column. It helps especially in the case of gathering different comments from different web pages.

4.5.2.2 Periodicity recognition

Some events may also happen more than once with some repetitive pattern, regularly or irregularly; these are the records that have the same value attributes for their name, category and subcategory but are different in the other attribute values. In addition, some webpages do not provide any relevant information on events patterns. These two aforementioned issues were solved through periodicity recognition. The information from periodicity recognition function is stored in column named “how_often”.

An event is regarded as regular, if records with the same value for “name”, “category” and “sub-category” columns repeat at least three times in the local database. To calculate the “how often” information, the column “startdate” is used. We calculate the average time interval between two dates, and then the periodicity is matched. The periods are flexible in the matching rules. Indeed, due to various reasons such as weather conditions, holidays, and organizational regulations, a regular event may be held in an earlier or a later date. Therefore a weekly event may not happen exactly every 7 days. It is the same for the recognition of monthly and yearly events.

4.5.3 Synchronization

The synchronization between the local database and the OSM event database is realized by a SQL transaction. The synchronization consists of three steps; extracting the changes, migrating the changes to the database, and removing the respective tiles from cache. Table 4-1 shows a sample type of data saved to the database. The data quality is checked manually after adding to the database.

id	name	startdate	enddate	category	subcategory	url	h
361	EHC Red Bull München – Hamburg...	2013-12-13	2013-12-13	Concert	Concert	http://www.m...	
362	SC Riessersee – Bietigheim Steelers	2013-12-13	2013-12-13	Concert	Concert	http://www.m...	
363	SC Riessersee – EV Landshut	2013-12-20	2013-12-20	Concert	Concert	http://www.m...	
364	EHC Red Bull München – Düsseldo...	2013-12-22	2013-12-22	Concert	Concert	http://www.m...	
365	EHC Red Bull München – Augsbur...	2013-12-26	2013-12-26	Concert	Concert	http://www.m...	
366	Asbjørn	2013-12-01	2013-12-01	Concert	Rock & Pop	http://www.m...	
367	Hiatus Kaiyote	2013-12-01	2013-12-01	Concert	Rock & Pop	http://www.m...	
368	Christina Stürmer	2013-12-02	2013-12-02	Concert	Rock & Pop	http://www.m...	
369	Velojet	2013-12-02	2013-12-02	Concert	Rock & Pop	http://www.m...	
370	Jake Bugg	2013-12-03	2013-12-03	Concert	Rock & Pop	http://www.m...	
371	Mount Kimbie	2013-12-03	2013-12-03	Concert	Rock & Pop	http://www.m...	
372	The bianca Story	2013-12-03	2013-12-03	Concert	Rock & Pop	http://www.m...	
373	Acoustic Fever	2013-12-04	2013-12-04	Concert	Rock & Pop	http://www.m...	

Table 4-1: Sample results stored in the database.

4.6 Capturing Volunteered Event Related Information through Social Media Stream Mapping

Although social media sites are like gold mines for social event detection, there are many challenges associated with these heterogeneous metadata. Furthermore, each event detection case requires a different clustering algorithm that should be defined using special techniques. There are possibly hundreds published approaches with different clustering algorithms towards the problematic of event detection in social media. Unfortunately, there is no clear common basis to be found in these works. Most of the works are targeted at a single specific platform. They are created independently from each other, leading to the lack of a common interface and process, which makes it hard to evaluate and compare different works against each other. This lack also hinders further research based on the existing works since researchers are forced to re-implement the works in order to evaluate and extend them. In addition, researchers face a huge overhead when dealing with the topic and are forced to replicate already existing implementations including the adaptation to different social media platforms and the presentation of the outputs.

Whereas a closer study of most social media platforms shows that they share many basic features, making it possible to use same algorithms and detection processes for all of these platforms. Various existing event-detection algorithms share a very simple architecture. Hence, to circumvent the aforementioned challenges, in this study a framework is developed that divides the general task of event detection in social media into independent subtasks and allows the development of platform-independent models. This enables the complete separation of data retrieval, event detection and presentation layers. The framework is built on a modular architecture, which makes the integration of any generic event-detection algorithm possible through a simple python adapter. The Modular structure of the framework also offers the possibility to integrate and adopt any clustering methodology easily. This general framework provides the possibility of event detection from any type of social media and has the potential to be used for event detection from multi social streaming sources. These abilities help the researchers to easily test, evaluate and compare different event-detection algorithms for different sources of information; here social media. Currently two data adapters for two different social media platforms (Flickr and Instagram) have been developed. Furthermore, the framework includes a web-based graphical user interface (GUI) that can be run even on restricted devices such as tablets or smartphones. The GUI has many interactive capabilities such as maps (Openstreetmap) and geo-spatial 3D plots for better visualization of the detected patterns. The outputs are also saved as comma-separated values (CSV) files, which makes further statistical analysis of the results possible in any statistical package. The framework requires various parameters for any online data stream such as location where events are searched (a bounding box), time of events (a period of time), parameters of clustering algorithm like number of contributions and common word usage. Upon running the tool, available data for the defined space and time are downloaded through the social media API and stored in a local event database. These data are then clustered in space and time based on initial parameters. The framework has been successfully tested on Flicker and Instagram platforms for different periods of time in different locations to detect many latent events. Due to the importance of spatial characteristics of the data in this work, only geo-tagged photos are stored.

When reviewing different works, it turns out that event detection is a challenging task in the field of data analysis in social media. The basic process comprises detection of meaningful patterns in available data, which can be used to infer social events which take place. In addition, researchers have to cope with a large overhead of processes not related to the actual problem but necessary for the overall task. Indeed before starting the actual core process several questions about the data acquisition and preparation arise. The programming interfaces of the platforms have to be examined and data retrieval and buffering have to be considered. Once the data has been made available it often requires additional preprocessing steps. It is only after these steps that the actual event detection algorithm can be considered, however there are still open tasks left in order to finalize the whole process. The implicit information inhered in the detected content clusters need to be extracted in order to describe the events; furthermore the results still should be presented and evaluated. Indeed aside from the actual basic detection process researchers face a huge overhead of additional tasks dragging away the attention from the actual problem and pacing down the process of research. Most of these additional tasks are shared across different implementations and are just replicated in new studies. Hence, this framework is developed to support the research community allowing them to concentrate on their actual task: the detection of events.

4.6.1 Software architecture

Several existing event detection algorithms and solutions have been examined and the most basic shared structure has been identified. These various event-detection algorithms share a very simple, yet powerful architecture. The generic structure of such algorithms can be seen in Figure 4-16. It is a linear processing pipe where the data undergoes successive transformation from raw data to a descriptive representation of the detected events.

The actual event-detection algorithm represents the core of this processing pipe and is surrounded by the data acquisition, the optional preprocessing, and the event construction. Event construction includes the extraction of inherent information from detected events (content clusters) as well as their representation. This structure emphasizes on the clear separation of different parts of an event-detection approach and encourages the creation of a modular framework, the independent development of different parts, as well as the reuse of already existing implementations. Using this model, a generic and extensible programming framework is developed to find meaningful patterns out of heterogeneous and unstructured online data streams. The separation of different processing stages comes with the benefit that each stage can be exchanged and replaced by an alternative implementation, allowing a dynamic adaptation to different objectives such as choosing a target social media platform or adapting the output of the process. However the development of such a generic framework requires some modifications to the previously shown processing pipe.

Indeed the event detection is not the only stage that requires preprocessing of the acquired data through data acquisition step; all stages in the process may require preprocessing of the data from the previous stage. Since the characteristics of data in different social media platforms differ from each other, preprocessing must not only be applied before the event detection step but an optional preprocessing/filtering step must be available for all stages. Figure 4-17 shows this generalized processing model. The basic processing steps stay the same as in Figure 4-16, but each of them is succeeded by an optional filtering step.

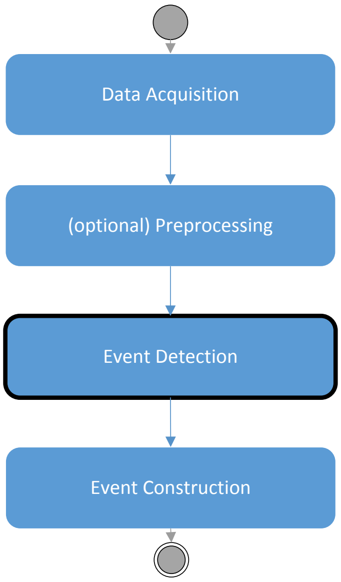

Figure 4-16: Basic event structure shared across common detection algorithms.

As can be seen, the new processing pipe is a linear combination of transformation and filtering steps and therefore an instance of the well-known pipes and filters pattern (Bushmann et al., 1996). The individual transformation steps have no cross-references on one another but solely rely on the data produced by the preceding step, granting the individual components a complete independence of each other as long as the data-contracts between them are met. This not only allows the individual components to be developed independently but also allows the construction of modular applications through a dynamic composition of data acquisition, event detection and event construction.

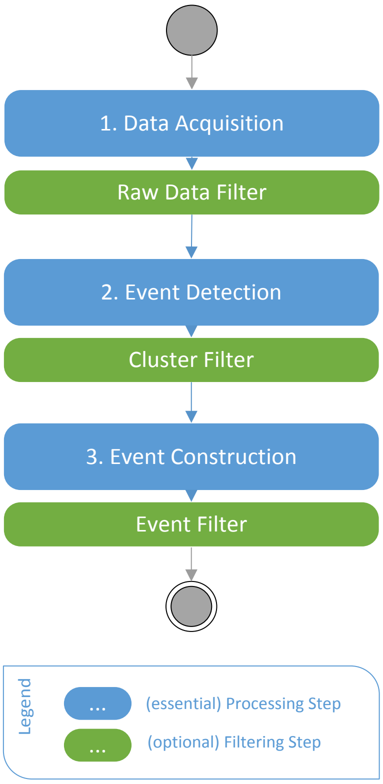

Figure 4-17: Generalized event-detection processing pipe.

As an additional enhancement, the framework comes with a graphical user interface easing the composition of different processing stages and providing an easy way of adapting its parameterization. This facilitates testing, evaluating and plain usage of the algorithms and also provides users with the possibility to get deeper into the matter by literally grasping the algorithms, comprehending the impact of its parameters and getting visual feedbacks. The framework also ships with an already implemented selection of all its components, including data adapters to Flickr and Instagram, predefined filtering steps, information extraction components and visualization features like maps, plots and charts. The concrete implementation of the framework and the graphical user interface is discussed in the following sub-sections.

4.6.2 Application framework

The framework is built upon the Python language which has recently received much attention in the academia; especially due to the clear structure of its code and many available standard toolkits. Python and similar languages such as R, have even become a de-facto standard in the field of data mining and machine learning. It is not only supported by field specific frameworks like NumPy (NumPy-Developers, 2014) and SciPy (SciPy Developers, 2014), which enable the language for mining big data; it is also equipped with numerous general third party tools (Oliphant, 2007). Furthermore, a graphical user interface based on a state-of-the-art combination of HTML and JavaScript is added to the framework. This separates the basic framework and the user interface in a server-client manner (Bushmann et al., 1996). It allows the basic framework to be the part that mines for events – using non trivial calculations on powerful machines – whereas the user interface can still be run on more restricted devices such as tablets and smartphones. This separation is more important when facing big data, since calculations can be outsourced to cloud, while the interactive part of the framework will have no constraint for client devices. In addition, since the uprising of web 2.0, HTML/JavaScript is prevalent in the field of light weight graphical interface design. Due to this combination and numerous available third party tools – ranging from simple data frameworks up to 3d graphic engines – developers are able to extend the framework without requiring knowledge in a Python specific graphical framework.

Model: The most essential part of the framework is the model representing the data flow in the event detection process. This model specifies the data contract between different stages, therefore is the key in assuring a clean separation of the stages and providing the possibility for independent development of the different components (Figure 4-18). Like the event-detection process, it is also separated into three parts.

The contract for the data acquisition component is specified by the BasicContent model. It models the content acquired from an arbitrary social media platform and is kept very general only including parts which are shared by the great majority of existing platforms. Its representation contains a timestamp specifying the date of creation, a geo location reference and a text field. This model covers the ground case and is compatible to the majority of platforms, including Facebook, Twitter, Flickr and Instagram. But in order to support the generic nature of the framework the BasicContent model is supposed to represent the top most element of a hierarchy of content models. As can be seen in Figure 4-18, it can be extended to support any kind of content by a simple inheritance; such as AudioContent and VideoContent. The hierarchical model – using the concept of information hiding – allows the components to decide on the degree of information, which they want to extract from the models. Hence, it ensures the compatibility of existing components with future extensions while placing no restrictions on the extensions itself.

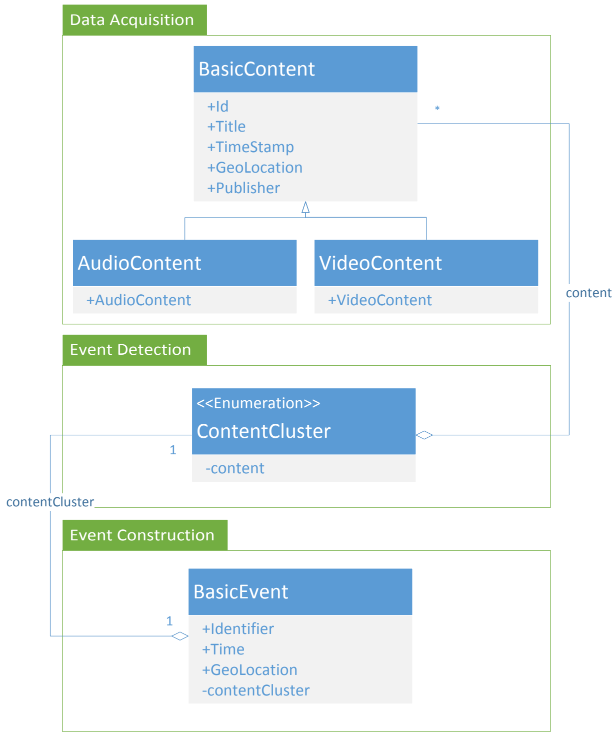

Figure 4-18: UML class diagram: Data Models.

The data contract for the second processing step (the event detection) is represented by the ContentCluster model. Just as with the content models, ContentCluster only covers the ground case and can be extended freely if the component needs to communicate any additional information.

The last data model belongs to the event construction stage and represents the final output of the whole process; the BasicEvent. This model represents a detected event and contains an identifier, usually representing the name of the event, as well as fields for the date and location of the event. It additionally holds a reference to the ContentCluster to which it belongs to grant access to the actual content related to that event. Again, the model can be seen as the most top element of a hierarchy of models, covering only the ground case and being open for any extension.

The discussed data models define contracts between different processing stages and represent the first step towards a modular system. But in order to ensure the independence of the sub-modules and in order to make the components substitutable, the exact interfaces of the individual processing steps still need to be defined. The model for these processing steps is displayed in Figure 4-19. Here again, the diagram is split into three parts representing the three basic stages of the previously introduced processing model. The interface for the data acquisition stage is defined by SocialMediaContentProvider which serves the sole purpose of retrieving all contents available on the platform for a queried timespan and a geo bounding box.

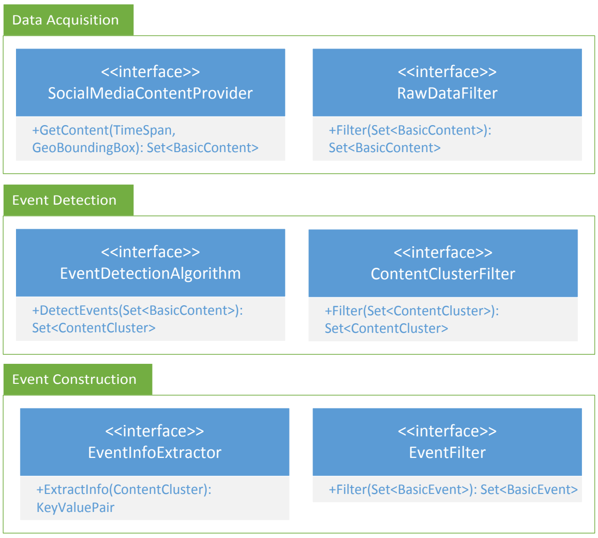

Figure 4-19: UML class diagram: Processes Interfaces.

As previously discussed, the separation and the intended independence of individual stages require an optional filtering step following each stage. This filtering serves the purpose of postprocessing the intermediate output, smoothing the data transition such as filtering outliers. The framework allows the application of multiple successive filters and the filter interfaces presented in Figure 4-19; RawDataFilter, ContentClusterFilter and EventFilter. The event-detection stage is represented by the EventDetectionAlgorithm interface to extract clusters from a set of contents to form events.

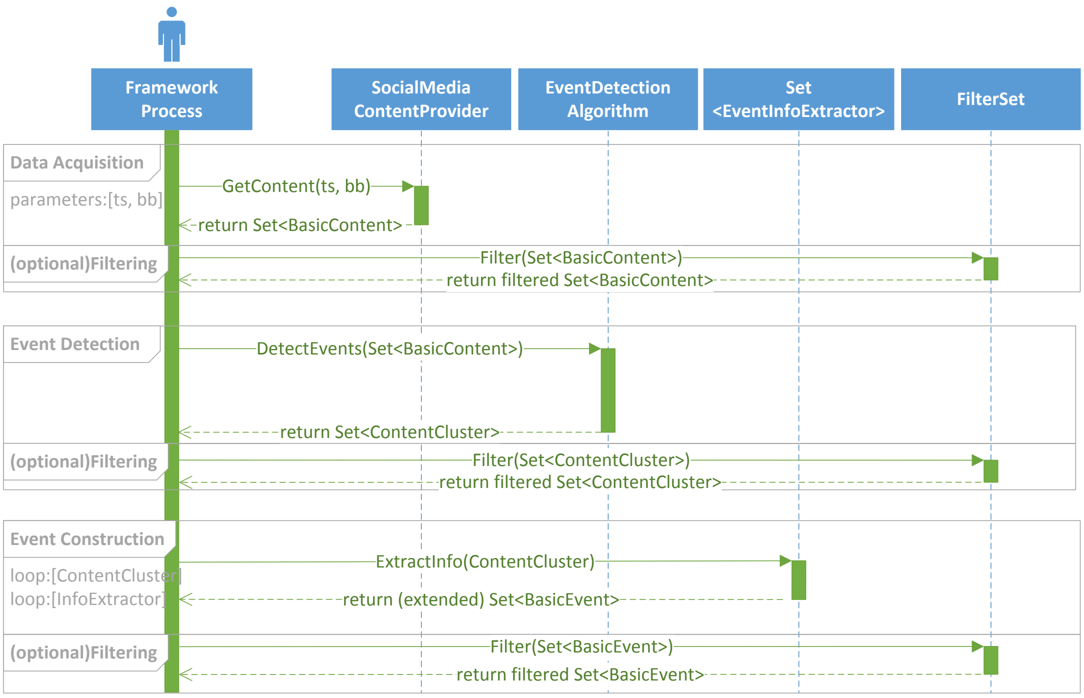

Figure 4-20: UML sequence diagram: general framework process.

The first two stages are represented with a similar implementation in the framework process. The process requires one data provider and one event detection algorithm. In contrast, the final stage (the event construction) is of a more dynamic nature. The construction transforms the raw content cluster identified as forming an event into a more descriptive representation embodied by the BasicEvent model. This model holds basic information about time, location and identification of the event and is ought to be extended depending on the purpose and objective of the actual application. Therefore, the construction stage is designed to allow a dynamic assembling of this representation, which allows defining a set of EventInfoExtractors to extract the desired information and dynamically extend the BasicEvent model. EventInfoExtractors may be used to infer time and location of an event but may also extract more abstract information such as statistics and graphical plots.

The framework process representing the actual interplay of the defined data models and process interfaces is shown in Figure 4-20. As can be seen, it reassembles the original process model of the Figure 4-17. Using the introduced models and interfaces, different parts of event-detection algorithms can now be developed independently.

Graphical User Interface: The graphical user interface enhances the framework by a graphical component and is targeted at the evaluation and production stage. The interface is built upon a combination of HTML/JavaScript and is to be run in a common web browser. It therefore is an optional component to the framework, separated from the core in terms of a client-server pattern (Bushmann et al., 1996). This not only allows the core to be outsourced to a powerful machine, but due to the widespread web browsers even allows restricted devices such as tablets or smartphones to be used as clients of the framework. The graphical user interface was developed in the manner of a drag-n-drop user interaction concept.

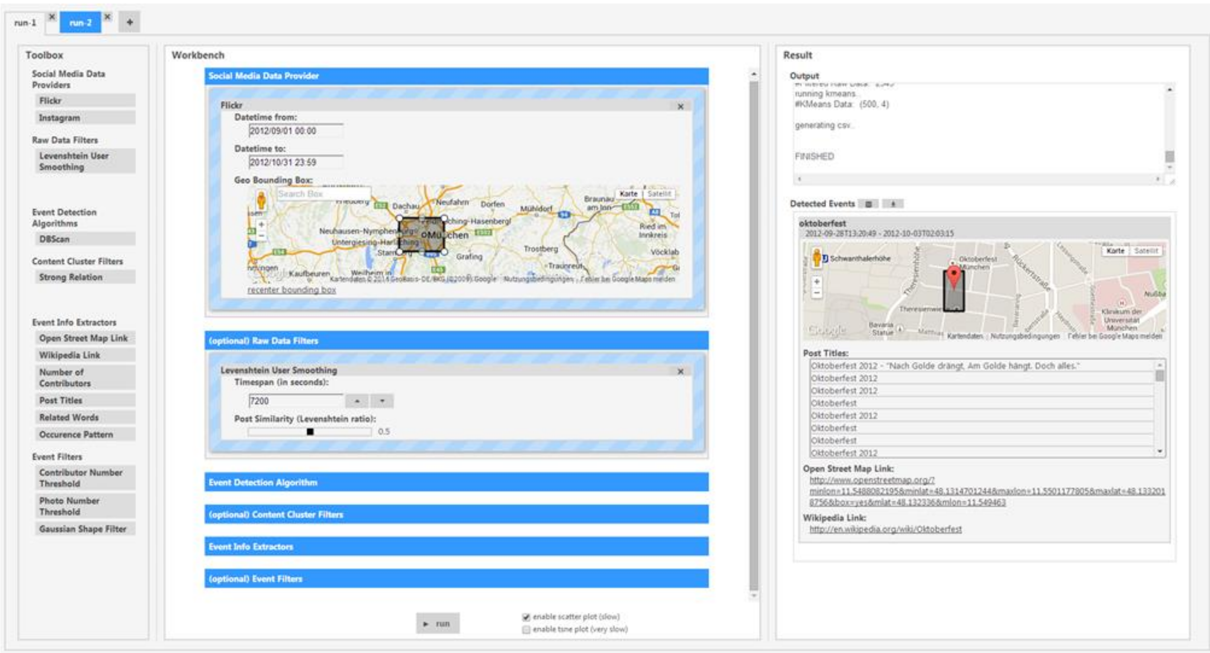

Figure 4-21: A screenshot of an exemplary configuration of the graphical user interface.

A screenshot of a basic and exemplary configuration of the interface can be seen in Figure 4-21. The view is split up into three main parts, the toolbox, the workbench and the result panel. The toolbox displays all available components as discussed in the previous sections. These components can be dragged into the workbench area in order to fill the drop-areas representing the different stages of the framework process. When being dropped, the components unfold and offer the possibility to configure their parameters. Any component compatible to the interfaces discussed in the previous sections can be integrated into the UI. It only requires the definition of a constructor and its parameters. For the unfolded workbench that represents a component, the widgets can be drawn from a set of predefined HTML/JavaScript widgets. The screenshot in Figure 4-21 contains a small exemplary subset of them; a map, a date-time chooser as well a slider widget. Due to the nature of HTML/JavaScript there is no restriction on extending this set. A great majority of the widgets is provided by jQuery-UI (Foundation, 2014) but the user is free to use any additional third party libraries. With this concept the user can integrate its components into the user interface by simply providing a constructor, its parameters and its visual representation in a simple configuration file.

After composing the components and adapting the parameters, the result panel shows a live output of the original code while the algorithm is running, allowing easy debugging during runtime. After the algorithm is finished, the detected events also appear in the result panel. The basic data of the events is displayed and a miniature map visualizes the approximate location as well as the approximate size of the events. Other results can be dynamically composed based on the actual BasicEvent implementation as well as their extensions from the user's choice of EventInfoExtractors.

For visualizing the results – like the workbench – it is possible to draw widgets from a set of predefined HTML/JavaScript elements. This enables developers to completely adapt the visual output of application. The predefined components include different visualization elements such as maps, simple hypertext links. In addition, the graphical interface offers a selection of plots, including a 3d scatter plot as well as a state-of-the-art t-Distributed Stochastic Neighbor Embedding (t-SNE) plot (Van Der Maaten and Hinton, 2008) for visualizing results. Output plots are used to evaluate the impact of different parameter settings on the whole data set. The graphical user interface also exports the result data in a comma separated value (CSV) format reflecting the BasicEvent model. The framework is publicly available and can be downloaded from https://polous@bitbucket.org/polous/smm.git.

4.6.3 Framework outputs

A simple clustering-based method was used as an exemplary event-detection algorithm in order to evaluate the flexibility of the developed framework.

The assumption is that during the happening of an event people publish an unusual concentration of content related to this event. With this assumption detecting an event comes down to separating the content belonging to such a concentrated cluster from the rest of the data. This is done by detecting clusters of dense data regarding temporal, locational and textual features in the social media content. We use here a Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996) as an exemplary event-detection algorithm in the framework. Due to the simple interfaces of the framework, the DBSCAN algorithm was integrated using a simple Python adapter. After this integration, the framework is ready for new setups; choosing a target social media platform, applying preprocessing/filters, and defining the information which should be extracted from the events.

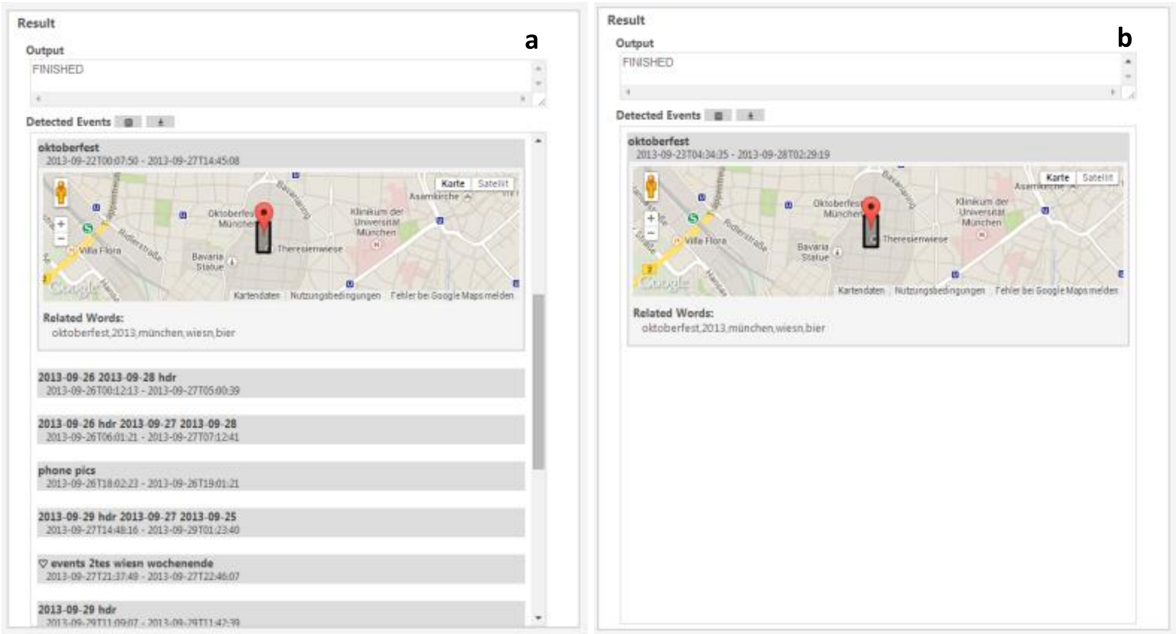

Figure 4-22: a) Result of a simple Event-Detection-Algorithm on Flickr, October 2013, Munich, b) Result of a simple Event-Detection-Algorithm on Flickr, October 2013, Munich, refined with the filters.

As a first step, the algorithm was targeted at the Flickr and Instagram adapters to provide access to the content of these platforms with a temporal SQLite (SQLite, 2014) database as a cache in between. The results for the application of the algorithm to data published at Flickr around October 2013 in Munich are shown in Figure 4-22.

As can be seen in Figure 4-22, some events (17) were detected, one of which is “Oktoberfest”. Oktoberfest is an annual Bavarian fair held around the beginning of October and located at the Theresienwiese in Munich. The algorithm detected its existence solely based on the content that users published around this date. This information was then used by the framework in order to extract knowledge from the content of the detected clusters such as a name, the presumable time and the location of the event. The extracted knowledge used for the purpose of evaluation includes; the basic event description features and an additional list of words likely related to that event. The visualization of the event and the extracted information can be seen at the bottom of Figure 4-22 where the map widget indicates the event’s presumed location. It approximately matches the actual location, which is the Theresienwiese in Munich. The related words list also seems to match quite well, including words like “munich”, “2013” and “beer”.

Although the algorithm was accurate with the “Oktoberfest” event, it also produced some false positive detections, suggesting events consisting of random words, like “2013-09-26 hdr 2013-09-27 2013-09-28” event. This probably happened due to the structure of the clustering algorithm, which only applies a density-based method. The result is refined and improved without adapting the algorithm itself, by simply using additional filtering and preprocessing steps. The results of an adapted run using filters such as user smoothing, thresholds, and matching the shape of patterns against a Gaussian distribution, is shown in Figure 4-22. As can be seen, the random events disappeared and the result was improved without changing the algorithm.

In addition to the filter extension, the framework also offers the possibility to retarget event-detection algorithms to other social media platforms. For instance, exactly the same configuration of the above setup was retargeted at the Instagram platform. This was done by simply selecting the Instagram provider from toolbox in the user interface and dragging it onto the social media stage in the workbench. The result of a run targeted at the same area and timespan is demonstrated in Figure 4-23. The “Oktoberfest” event was detected again, but this time from content published at Instagram. These detected events are manually checked before storing in OSM-event database.

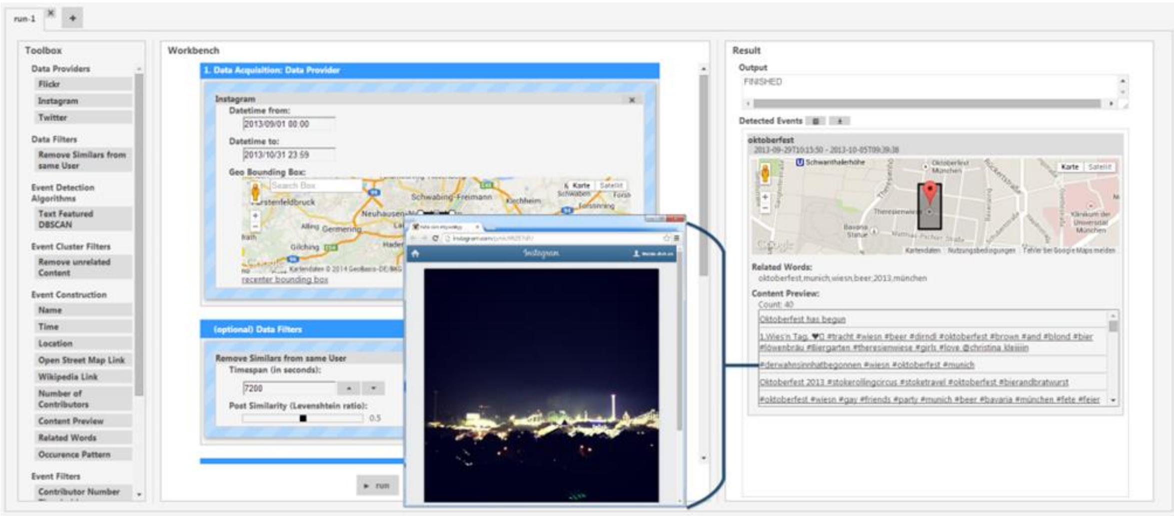

Figure 4-23: Result of a simple Event-Detection-Algorithm on Instagram, October 2013, Munich, refined with Filters.

5 UTILIZATION OF VOLUNTEERED GEOGRAPHIC INFORMATION

This chapter presents different mechanisms that have been developed to utilize, further enrich and analyze collected event information by contributors or through the web crawler (Figure 5-1). These mechanisms include: a) an android application which is discussed in chapter 4, b) a web application to provide spatial service to users, and c) a spatio-temporal analysis and visualization application to analyze gathered data and facilitate the engagement of volunteers to higher levels of involvement than merely data collector. The android application has been discussed in chapter 4, therefore it won't be covered in this chapter. The system architecture and technical details of other mechanisms are presented in this section.

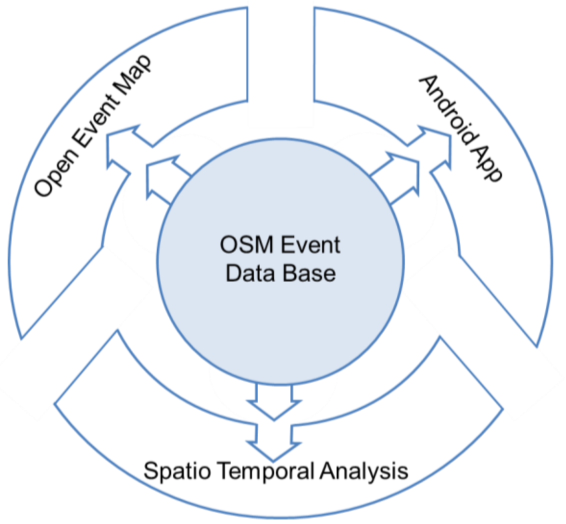

Figure 5-1: Three mechanisms to utilize the OSM event database.

The collected data in the locally deployed event-OSM database are first converted into easily searchable SQL database format. The relational algebra of SQL can offer significant reasoning ability for managing information in a database. Many operations are possible by holding combinations of spatial, temporal and semantic queries. Cron – a time-based job scheduler in Unix-like computer operating systems – runs a special script in the server every minute to convert event data into SQL database format. This script generates osmchange files in .osc format which contains addition, modification or deletion of any OSM Element within the last minute. Each .osc file is processed by OpenEventMap script and inserted into a new table in the local server called search_event. During this process, multiple key-value pairs related to a single event are converted into a single row in the new table format (Figure 5-2).

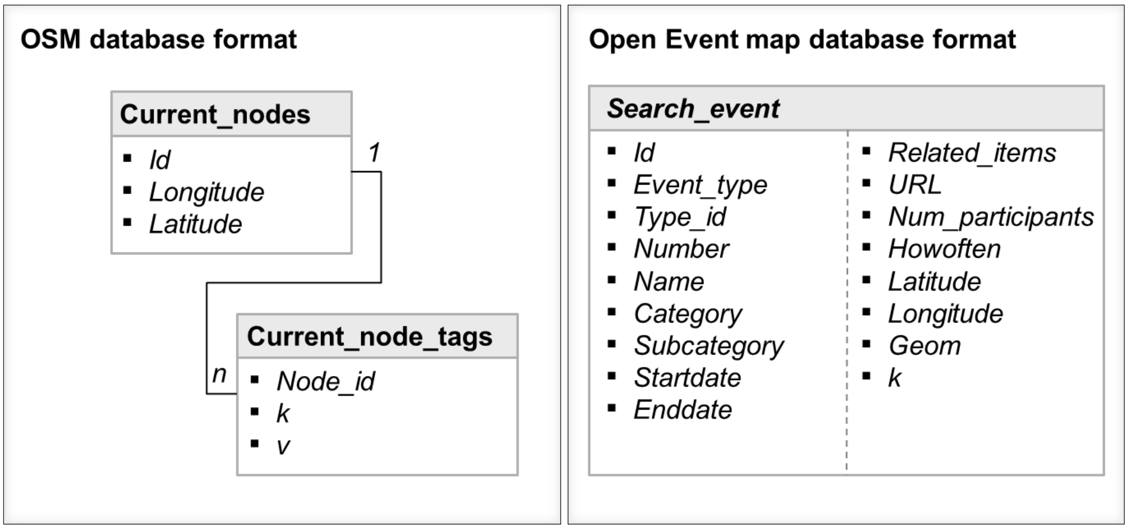

Figure 5-2: Comparison between OpenStreetMap and OpenEventMap table structure.

In the current event-OSM node-tags table, a single event data is stored across multiple rows as in Table 5-1.

node id	k	v
12345	event	yes
12345	event:0:name	Oktoberfest
12345	event:0:category	social
12345	...	...
12345	event:1:name	Frühlingsfest
12345	event:1:category	social

Table 5-1: Event-OSM table format.

Whereas two events are converted into two single rows in the search-event table (Table 5-2).

id	event type	type id	number	name	category	...
1	Node	12345	0	Oktoberfest	social	...
2	Node	12345	1	Frühlingsfest	social	...

Table 5-2: OpenEventMap table format.

The search-event table format makes it easier to do complex queries; searching more than one field at a time. For example the following query is possible: "SELECT * FROM search-event WHERE name LIKE %val% AND category='social' AND...". This query would be complicated to execute on current node-tags table.

5.1 Open Event Map

To utilize collected event information a "Location-Based Event Calendar" for Munich is developed. Location-Based Event Calendar provides geographic and semantic information, by answering the questions like "Where, When, and What". The information is saved in event-OSM data base and can be queried through a key/value pair in a web application named "OpenEventMap". The OpenEventMap implementation provides the possibility for end users to easily access geographic services. For visualization purposes, semantic and temporal information along with their spatial coordinates are extracted to be displayed on map (Figure 5-3). Results are color coded based on types.

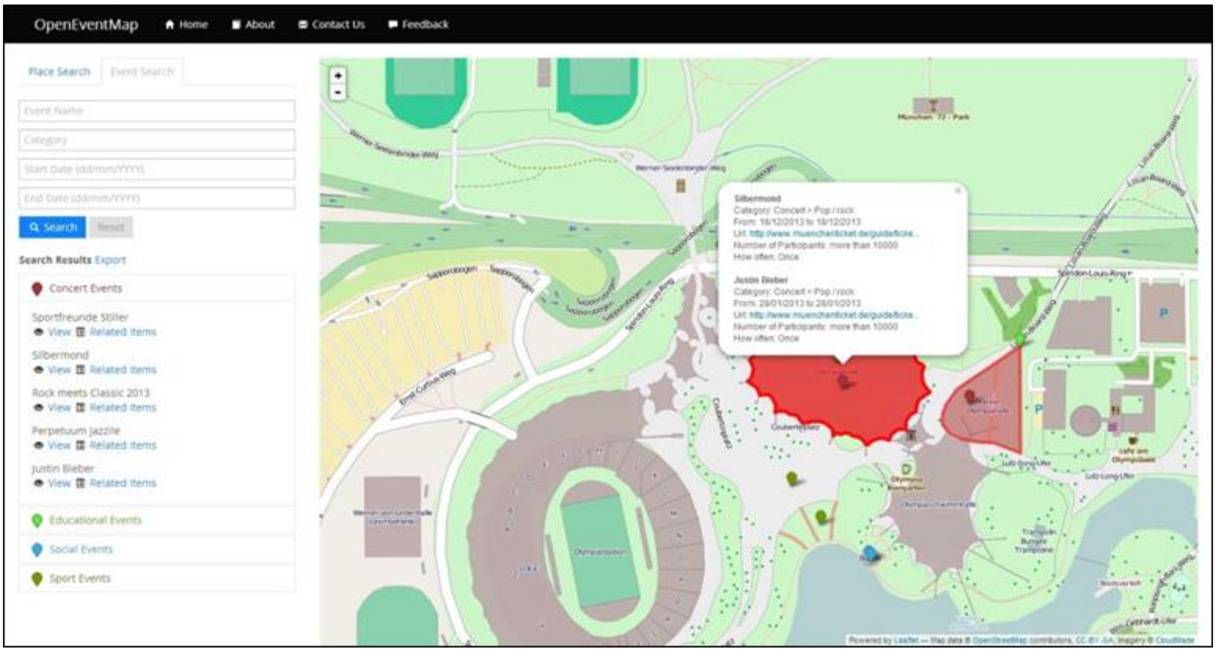

Figure 5-3: Web application screenshot (www.openeventmap.tum.de). Involving objects of the selected event are highlighted in red.

Four parameters can be used to filter the search results: event name, category, start date, and end date. For event name, category and sub category parameters, the exact match in the database is searched. But start date and end date create a date range and return all events that fall under a given range. Date filter also takes into account the repeating events. For example, an event that occurred in March 5, 2012 and has a yearly pattern will be found when searching for events between March 1, 2013 and March 10, 2013. If search is performed with none of the parameters entered, it returns all matching events within given map region such as in Figure 5-4. This tool also provides the possibility of searching places by name through OSM Nominatim API[9].

[9] http://nominatim.openstreetmap.org/

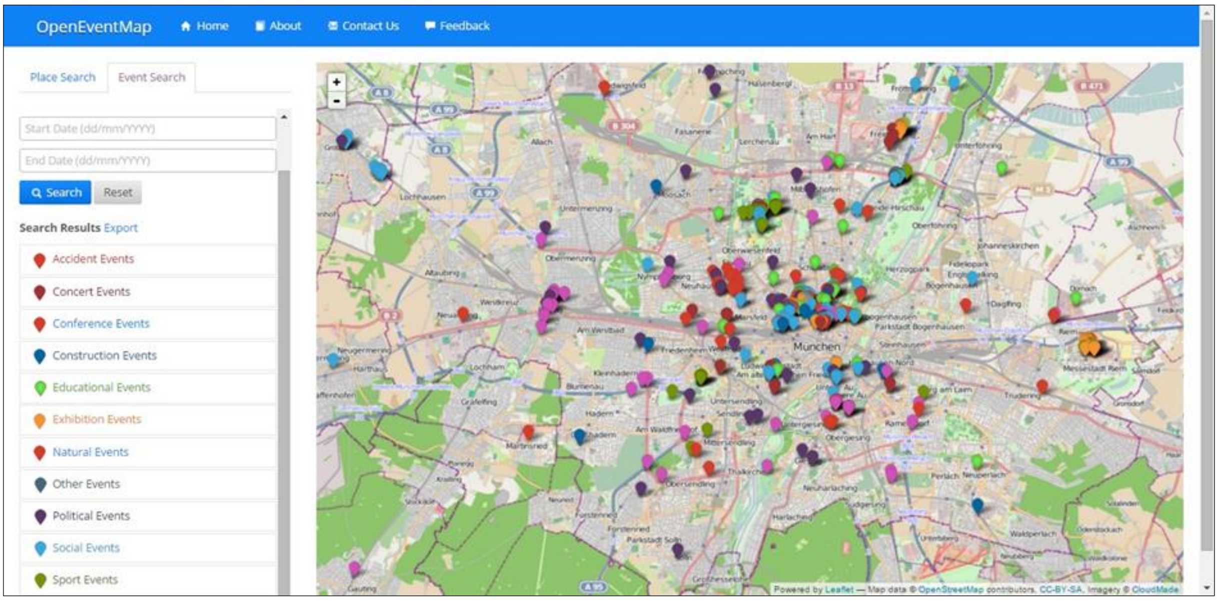

Figure 5-4: Searching all events in a bounding box.

5.2 From Volunteered Geographic Information collection to synthesis

As discussed in the previous sections, there are different challenges in using VGI data such as redundancy of gathered data as they have been generated by amateur users. The redundancy of the data comes from the fact that multiple contributors may gather similar information in parallel; especially when they use different terminologies and have different perceptions of a phenomenon in space. Heterogeneous user-generated information can substantially restrict usability of the data for decision-makers and especially citizens (Carver, 2003; Sieber, 2006). Although, expert users and researchers have access to sophisticated analysis and visualization tools that can help them to discover meaningful patterns by exploring complex and multifaceted VGI content, this is not the case for most of the amateur users. Furthermore, current online mapping developments have actually not yet managed to handle most kinds of VGI which contain qualitative and vague information. Indeed, when it comes to deal with VGI specificities, little work has been done (Deparday, 2010).

To improve the nature of the data, applications and tools should be designed in a way that can prevent contributors from using different terminologies and at the same time structure the gathered data (like event tag). Indeed our event-gathering mechanisms guarantee a much higher quality for spatiotemporal information comparing to previously available approaches. Second, to assist novice and average users with special technical constraints to operate on web, visual-analytic tools should be adopted and redesigned; like accessible web-based tools for visualizing results.

This sub-section is dedicated to a visualization and analysis tool for volunteers and professional users to analyze crowed sourced event data, which aims to facilitate the engagement of volunteers in the concept of VGI to much higher levels than a mere data collector. Recent advances in sensor hardware and software design and production has enabled the usage of multiple low cost electronic sensors almost in every device. Although this may dissolve to some extent the need of volunteer's engagement as human sensors, it critically raises the need for higher involvement of intelligent volunteers in analyzing and interpreting an explosive growth of available digital data. Indeed, the amount of stored data is multiplied rapidly in a course of few years. This shows the urgency in shifting the main focus of VGI from collection to analysis and synthesis.

Effective visualization and analysis capabilities have always played a key role in achieving concrete conclusions and interpretations from data, without which identifying the usefulness of data for users is impossible. Event Visualization and Analysis Tool (EVT) is a web based application[10] developed for the visualization and analysis of events and their attributes in the event database, however some of the features of the application can be used by importing external data. MapQuest has been used as the base map to visualize events and their features on it. The main focus of the EVT application is to provide users with a one stop solution for visualization and analysis needs. In addition, user friendliness and ease of use of the tool was designed with consideration of the fact that untrained volunteers are a main user group who may want to use this tool. To reach the maximum number of users a web based application is developed that enables using the tool on any internet enabled device which had a browser installed on it. To follow the spirit of open source concept, all used services and tools for developing the EVT are open sources.

The application supports two types of visualizations; map-based visualization and standalone analysis. Map-based visualization refers to visual output which has a map in the background as location reference. The visual output can be in the form of markers, heat map or pie charts on top of the map depending on the functionality being used by the user. Standalone analysis refers to other visual outputs of the analysis including charts and graphs which don’t have any map as the background.

5.2.1 Application architecture

Event Visualization and Analysis Tool (EVT) is a web-based application with the client-server architecture. The only requirement for a user is to have an internet enabled device with a web browser.

The general flow of information begins with the client accessing the application via web browser. The high-level application architecture diagram can be seen below in Figure 5-5. Details of the architecture and the information flow are discussed in the subsequent sections.

[10] http://www.openeventmap.tum.de:8080/EVT/

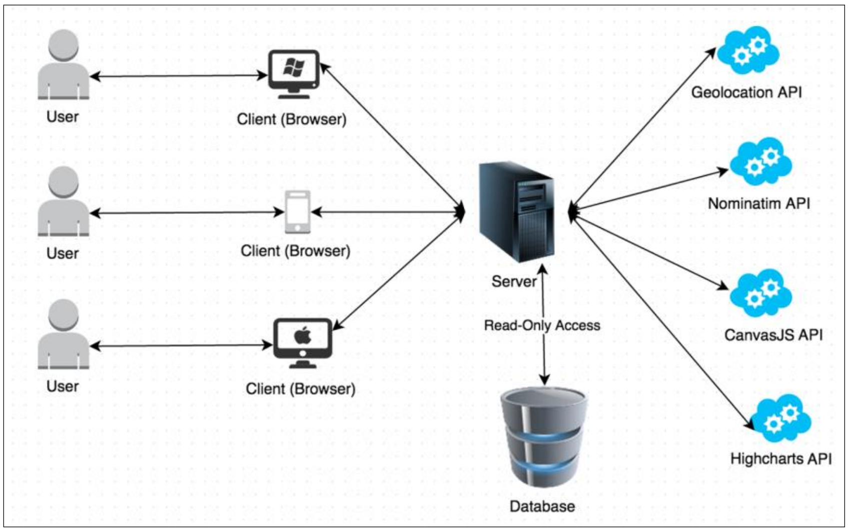

Figure 5-5: Overview of the application architecture.

EVT has been developed using Java development kit 7 and Java Enterprise Edition. As Java is open source and can be run on any hardware and software platform. In addition, it has a strong ecosystem that provides extensive up-to-date APIs to build and develop enterprise applications.

5.2.2 EVT software design

EVT as a web application is developed using the MVC (Model, View, and Controller) software design architecture. MVC ensures separation of concerns and issues between different components and modules of the software to manage complexity in the software code. Figure 5-6 shows the MVC-based software architecture of the application and its technology stack. EVT has three layers with different types of components: presentation layer (View), business layer (Controller) and data access layer (Model).

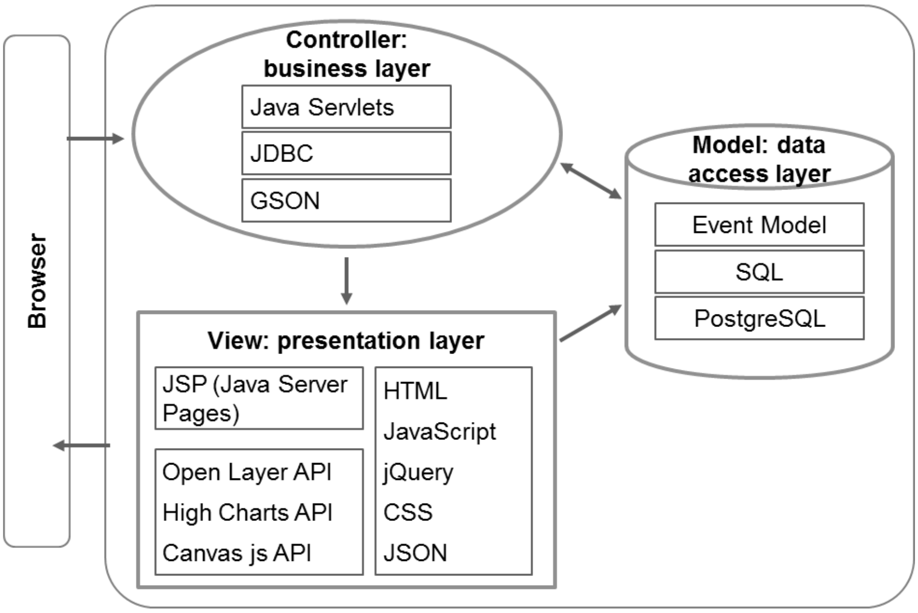

Figure 5-6: EVT Technology Stack and software architecture.

Model refers to data access layer of the application. The Java class EventInfoBean maps data from database to EVT controller. As EVT is a search and visualization tool, it only supports DML (Data Manipulation Language) queries on event database.

View refers to presentation layer of the application. The EVT application view is developed using Java Server Pages, JavaScript, jQuery and AJAX. User HttpRequest is wrapped in AJAX call to ensure the efficient usage of server resources. The EVT application pages are developed using jQuery and JavaScript. High charts[11] and CanvasJS API[12] are used for rendering charts in the EVT. The data from EVT application server is transferred to targeted client through JSON. EVT application uses Open layer API to render maps.

Controller refers to the business layer of the application. It is the communication point between view (Client) and model (database). Java Servlets API is used to develop the EVT controller. For every view of application, EVT has a separate Servlet controller, which takes user requests or queries and call OSM database. The controller prepares responses in JSON by using GSON API and sends it back to client.

5.2.3 EVT Deployment and functional flows

EVT application is deployed on Servlet Container "Apache Tomcat". Tomcat is a memory efficient, light weight and low configuration application server. The EVT application uses OSM database (OSM DB) and has access to other OSM-based services located on external OSM servers. Figure 5-7 illustrates the relations between the application server and other servers and services.

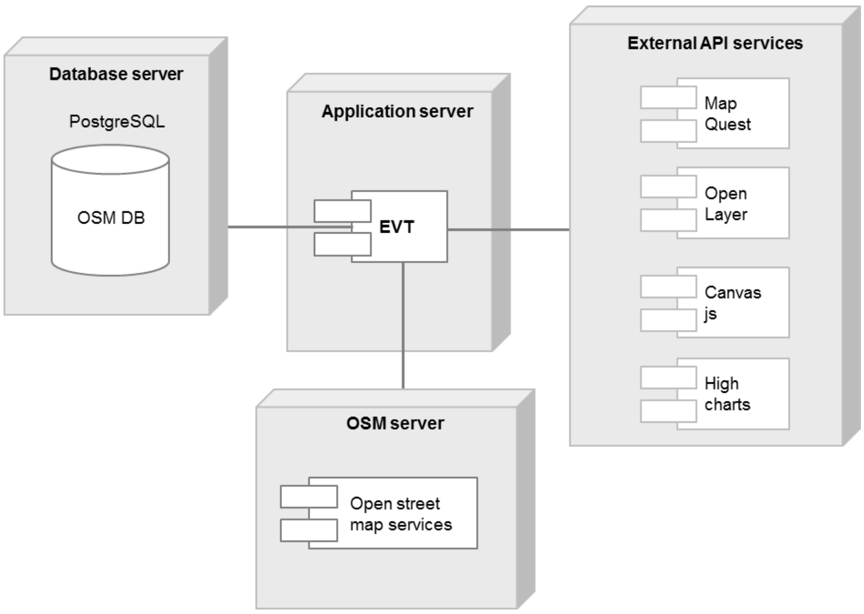

Figure 5-7: Services involved in the EVT application.

For the EVT application MapQuest[13] Open is selected as the third party source for generating tiles. It is a stable, reliable and free service available for the intended purpose. Furthermore,

[11] http://www.highcharts.com/
[12] http://canvasjs.com/
[13] http://wiki.openstreetmap.org/wiki/MapQuest#MapQuest-hosted_map_tiles

OpenLayers[14] has been used to visualize maps, which is an open source JavaScript library. OpenLayers is easy to implement and there are multiple features available as plugins for it.

5.2.4 Application Features

5.2.4.1 Spatio-temporal queries feature

The sequence diagram in Figure 5-8 shows the functional flow of advanced search event using external geo location services. It starts with a query searching an event, based on different criteria such as current location, name of place or a date. The current location of user and place names can be retrieved by external geo location services or can be entered manually depending on the way that client triggers event search. This method wraps the input parameters in HttpRequest to Search Event controller of the application. The controller will construct the search query based on input parameters. After accessing the database, the controller processes the query to retrieve search results. The search results are returned from database Result Set interface to controller in Search Event Bean, where controller prepares JSON object of Search Event Bean and sends it back in the HttpResponse to client. Finally the client renders the results on map as markers on the places where the events will happen.

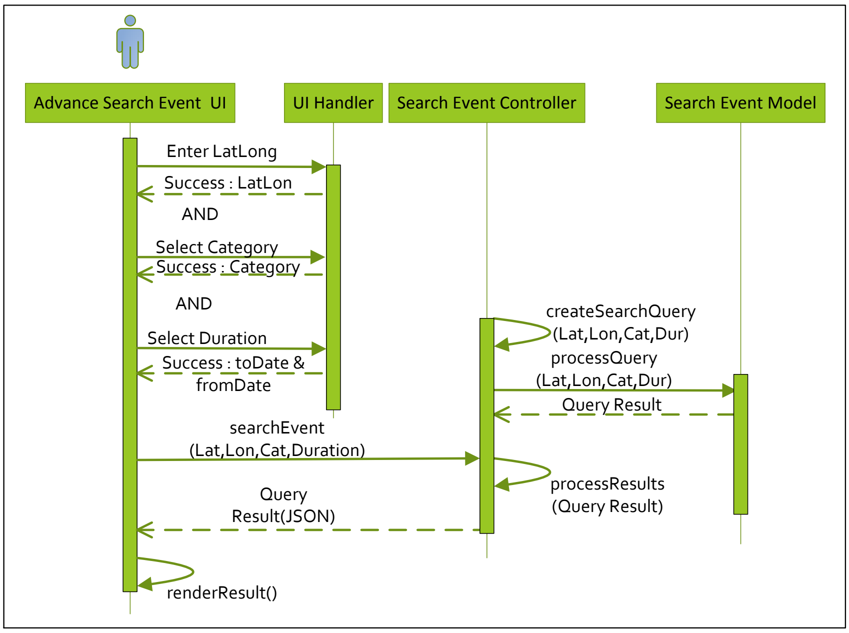

Figure 5-8: UML sequence diagram: advanced event search using external Geo-Location services.

The search feature can be used to find events in a configured radius from a particular place or a specific latitude and longitude combination. The search conditions are connected to each other using logical operations (AND, OR and NOT). The application uses a top-down approach for the execution of the features. The following filtering criteria can be defined as search conditions: latitude and longitude (place name, current location or bounding box), category, name, and duration.

The following scenario demonstrates how the spatio-temporal query features work and how users can mine the OSM-Event database simultaneously with the OSM-Object database to answer

[14] http://openlayers.org/

their queries. Assume a user travels to Munich and plans to stay in Hilton hotel between '01.09.2015' and '15.09.2015'; he would like to visit a social festival or a concert during his stay. He prefers a short distance venue (e.g. 10 KM) but in addition he looks for the events that are not between '07.09.2015' and '09.09.2015' (e.g. he has another plan for these days). Figure 5-9 shows the general framework for the aforementioned spatial queries. In the following, all involving steps are illustrated in more details.

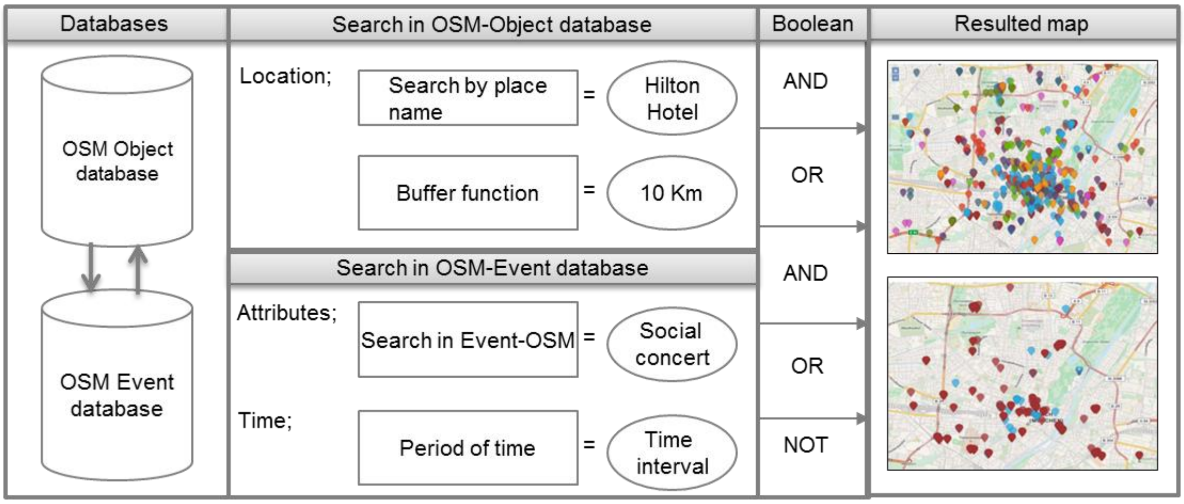

Figure 5-9: Spatial query in the databases.

Figure 5-10 demonstrates the result for searching Hilton Hotel and Figure 5-11 presents all events within a ‘10 KM’ radius from the ‘Hilton Hotel’ in Munich. It composes of two kind of analysis; search by place name and buffering. These queries are executed in OSM-Object database as well as OSM-Event database.

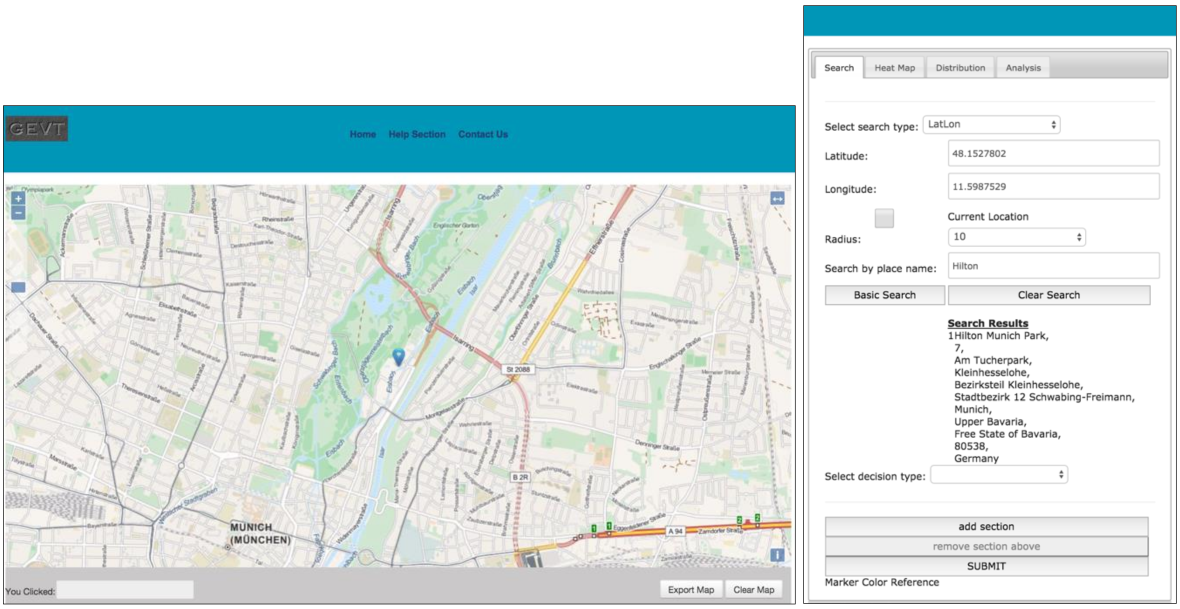

Figure 5-10: The map showing the location from 'Hilton Hotel'.

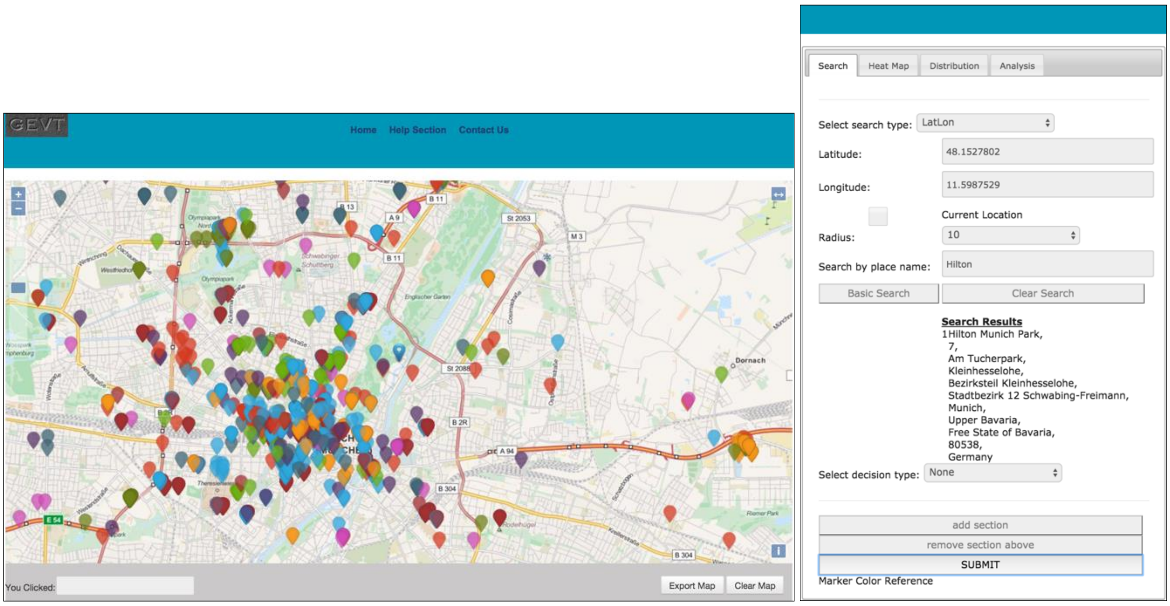

Figure 5-11: All events within 10 Km radius from 'Hilton Hotel'.

The next step is to preserve only the events that the user is interested in: social and concert. Figure 5-12 presents all social events available in Munich within 10 KM distance from Hilton hotel. All available social and concert events are illustrated in Figure 5-13. They were obtained through adding an AND operation on the previous snapshot. The quires are executed on the OSM-Event database.

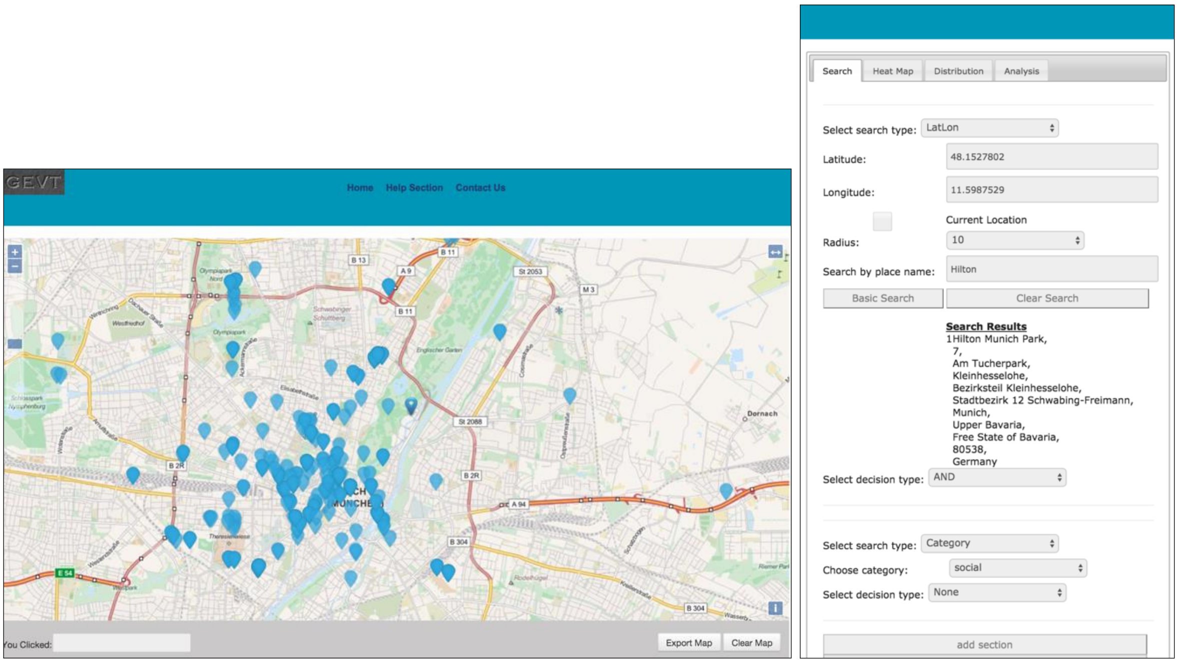

Figure 5-12: Social events within 10 Km radius from 'Hilton Hotel'.

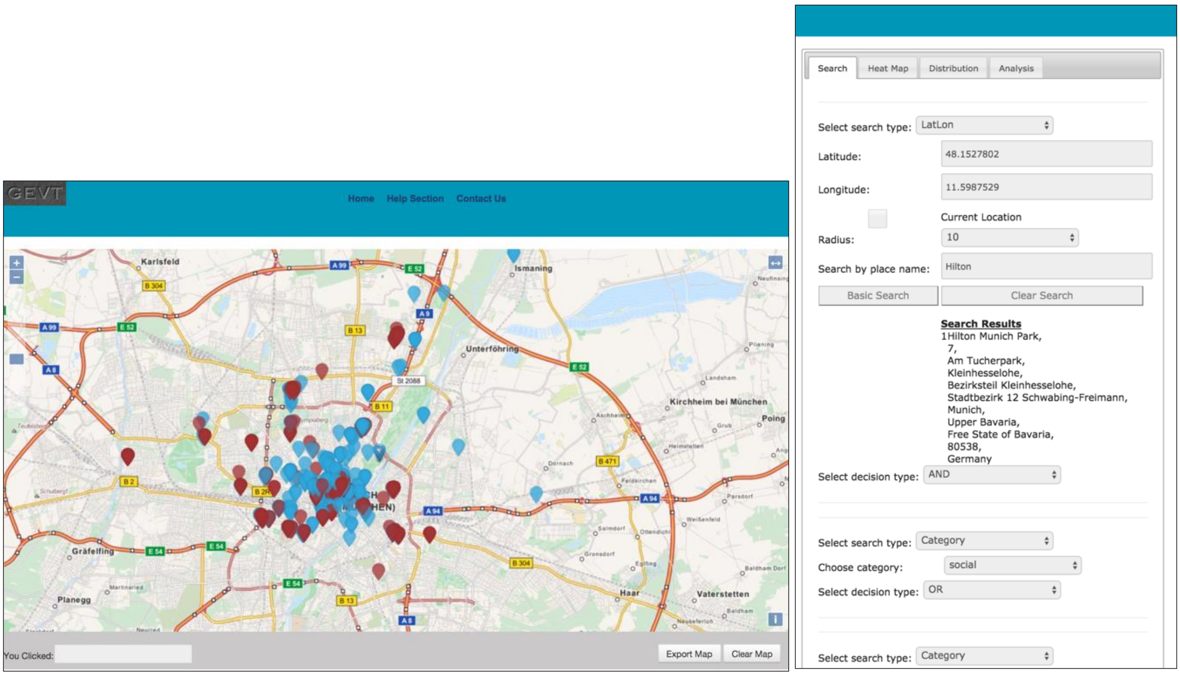

Figure 5-13: 'Social' AND 'Concert' events in 10 Km radius from 'Hilton Hotel'.

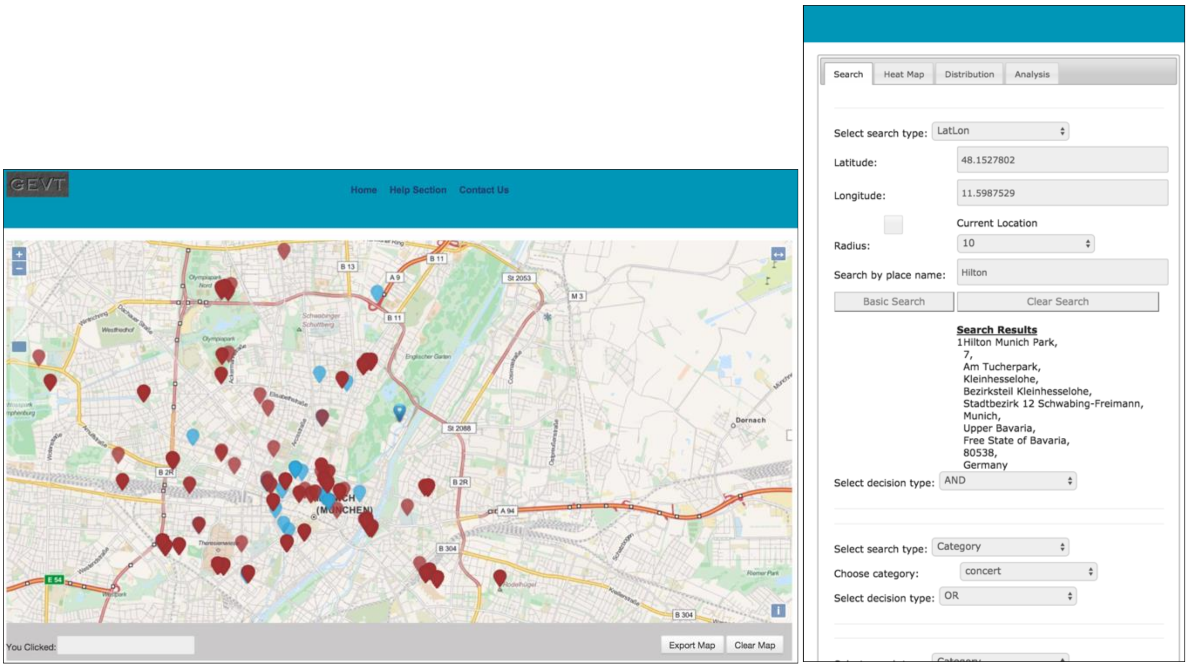

Figure 5-14: 'Social' AND 'Concert' events within 10 Km radius of Hilton Hotel between '01.09.2015' and '15.09.2015' but not between '07.09.2015' and '09.09.2015'.

The last part is to exclude all events that they don't meet the temporal criteria, this step is done by adding another search condition using OR and NOT operations respectively (Figure 5-14).

5.2.4.2 Map based visualization features

Aside from aforementioned spatio-temporal queries feature, the EVT provides two other map based analysis; heat map and distribution map.

The functional flow of heat map is illustrated in a sequence diagram in Figure 5-15. After triggering the Heat Map by client, the controller constructs the heat map data query based on input parameters. The search results are returned from database Result Set interface to controller in Event Bean, where controller prepares JSON object of Event Bean and sends it back in the HttpResponse to the client. Finally the client renders the heat map on map, using open layer API.

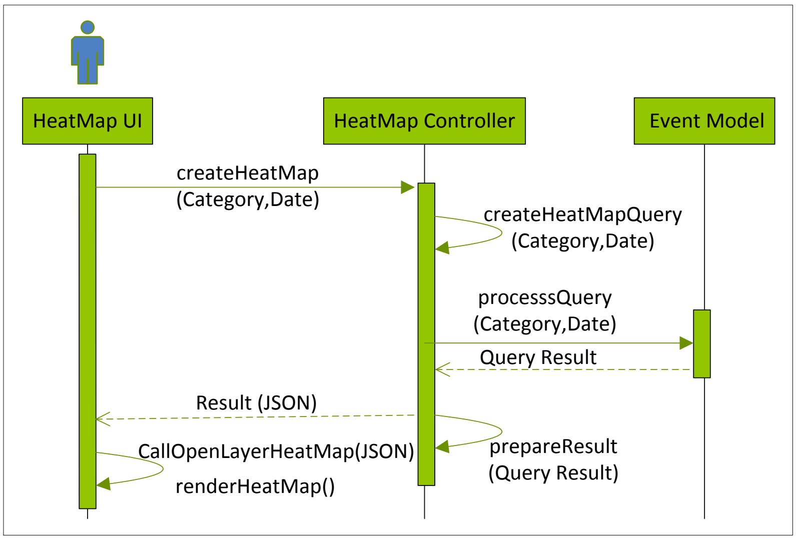

Figure 5-15: UML sequence diagram of heat map.

The filtering criteria for creating heat maps can be category, subcategory, time window, or any combination of them. Multiple categories subcategories can be considered for the selection of events. Events which fulfill all specified criteria are eventually used for creating the heat map. Figure 5-16 shows the heat map of Munich for selected event types in a specific time frame (02.08.2013 to 31.08.2015).

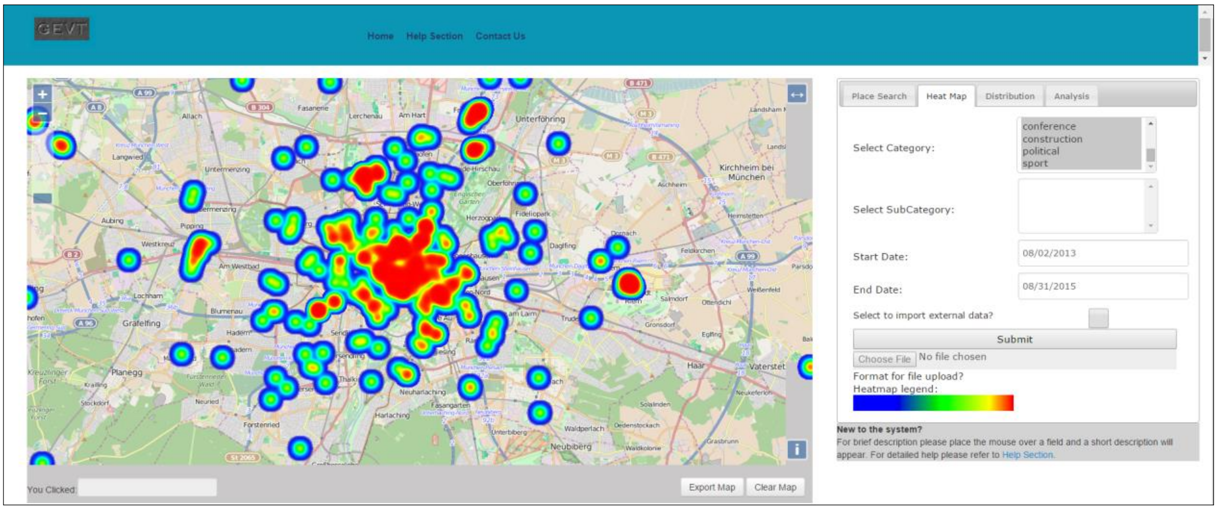

Figure 5-16: Heat Map of Munich showing selected events.

As mentioned before another map based visualization feature which the EVT is capable to generate is distribution feature. Figure 5-17 illustrates the functional flow of distribution feature for events according to category on map. It follows the same principal as the previous sequence diagrams.

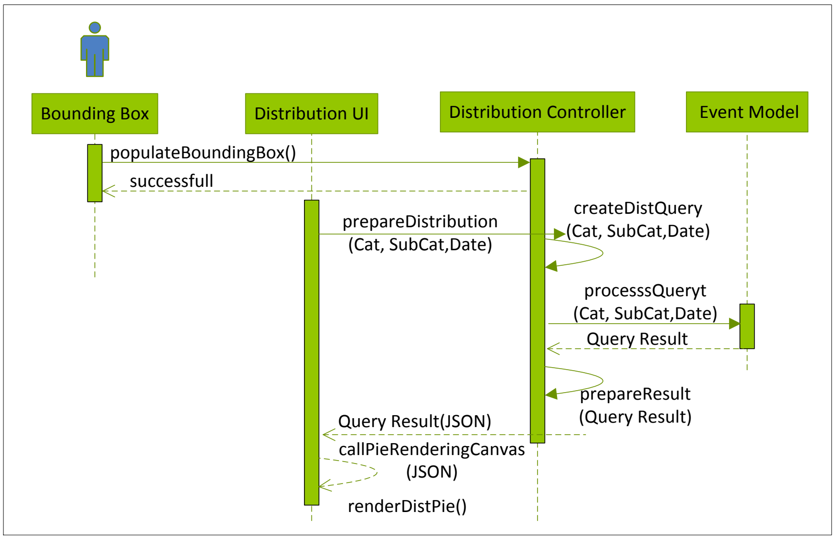

Figure 5-17: UML sequence diagram: Distribution.

The distribution feature represents statistical distribution of information in the form of pie charts on the map. The value used for the statistical distribution is the number of events which fulfill selected criteria. The size of pie charts is proportional to the number of events. The sectors in the pie chart correspond to distinct categories. In addition to the available database, the users also have the flexibility to upload external data for further analysis. The filtering criteria to generate pie charts can be a combination of: bounding box, time window, category, and subcategory. Figure 5-18 shows the distribution map of Munich for selected event types in a specific period of time.

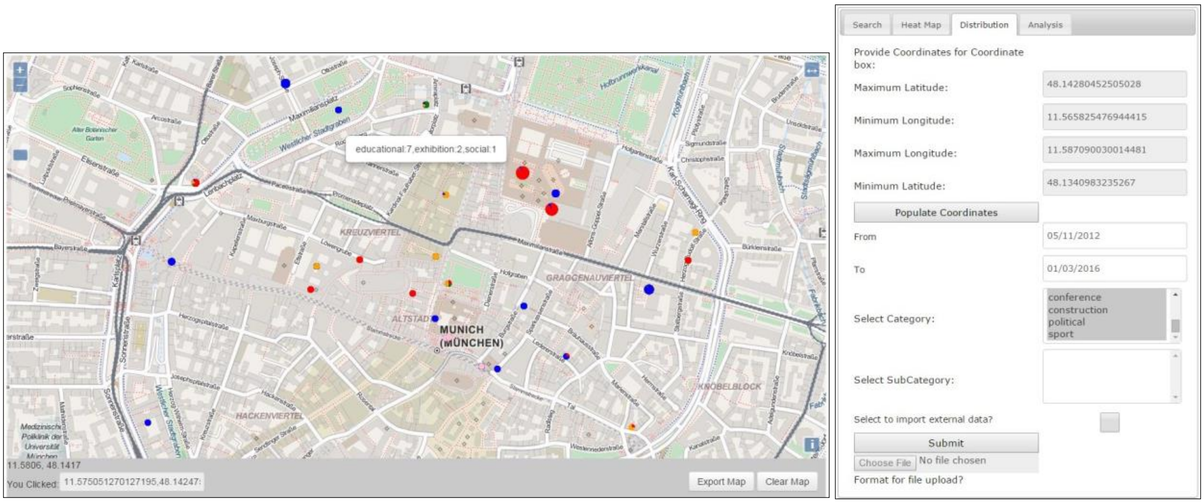

Figure 5-18: Distribution map.

5.2.4.3 Standalone analysis feature

Figure 5-19 shows the functional flow of standalone analysis features used to investigate event-OSM database. To create a chart, user should select chart type (such as a bar chart) and axis information (such as category and number of events for x and y axis). As can be seen in the sequence diagram user should first define the bounding box. The information follows in the same principle discussed already but client renders charts using High charts, canvas JS API and Protoviz[15] on client side to visualize the results.

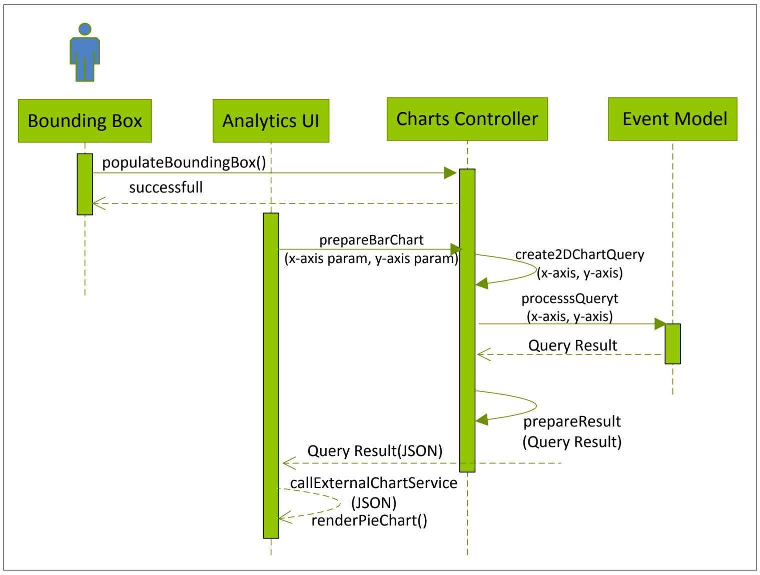

Figure 5-19: UML sequence diagram: 2D bar charts.

The analysis feature uses analysis and visualization techniques to view events based on several selection criteria such as time, space, category, and number of participants. EVT analysis features provide multiple options for event data visualization. In the following, different scenarios were designed to show the EVT capabilities in analyzing and investigation of collected event data. The event information collected either by volunteers or through a web crawler and might not reflect the real situation of Munich.

Figure 5-20 investigates the database in more detail and demonstrates the number of events in each district in Munich in 2012.

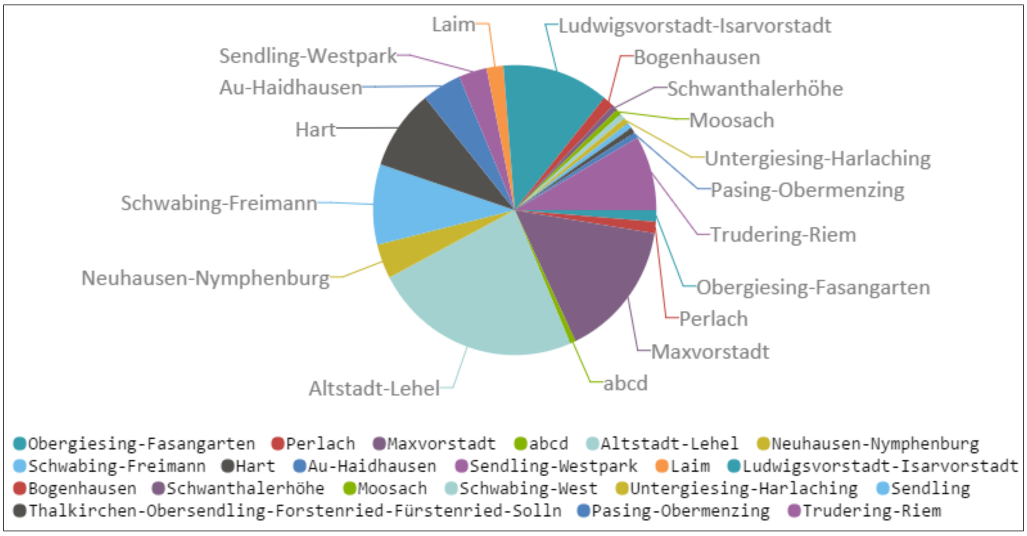

Figure 5-20: number of events per district in Munich.

[15] http://mbostock.github.io/protovis/

As can be seen, majority of the events are happening in Altstadt-Lehel, Maxvorstadt, and Ludwigwigsvorstadt-Isarvorstadt districts respectively.

A user might be interested in temporal distribution of events at a specific spatial scale (region or bounding box), e.g. number of events in different seasons in a year. Figure 5-21 shows the distribution of events in different seasons of 2012 in Maxvorstadt district.

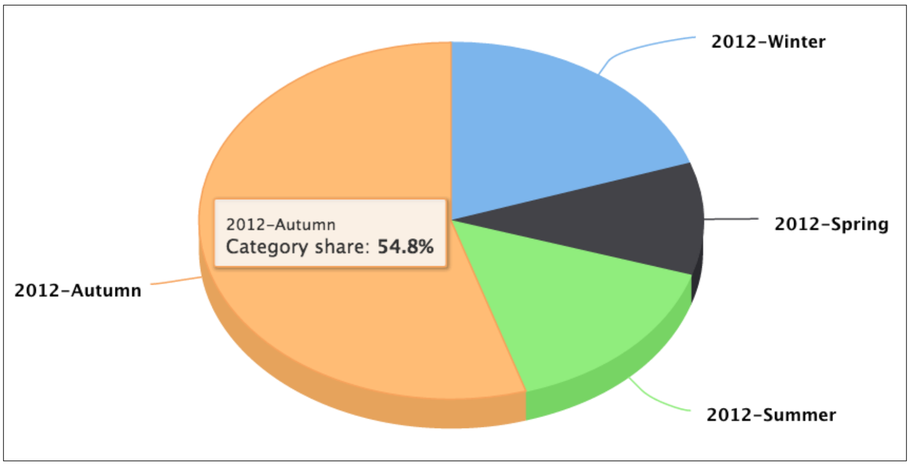

Figure 5-21: Distribution of events in different seasons.

In addition, a user can explore changes in the number of events for different categories at different time resolution for a specific bounding box or district. Figure 5-22 explores these changes for seven categories; accident, educational, exhibition, political, traffic and party between 2011 and 2015 in Maxvorstadt district.

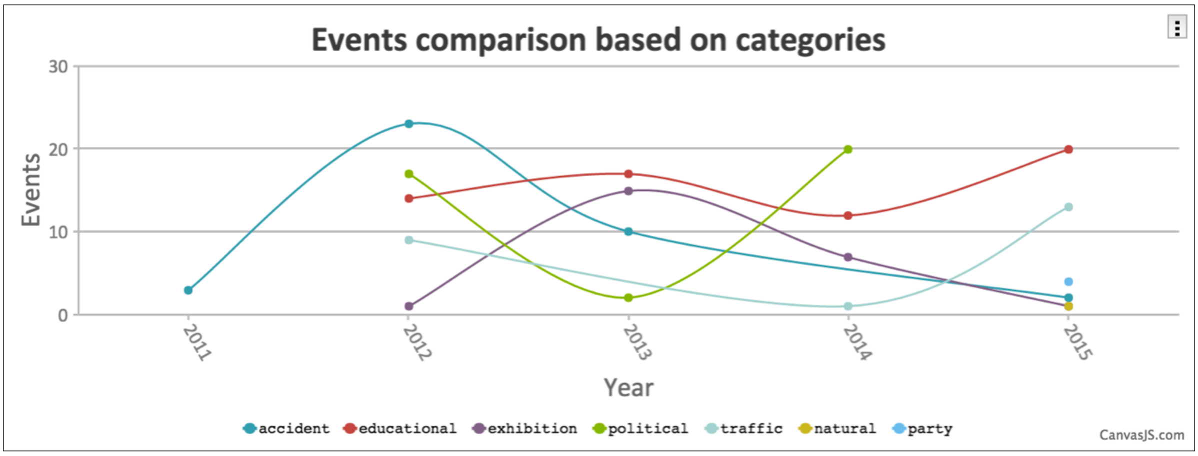

Figure 5-22: Trend of changes in the number of mapped events (crawled events are not included).

Another scenario would be to investigate the database for Munich. Figure 5-23 shows the number of events mapped from 2012 to 2015 in Munich. The rapid increase in the number of events in 2015 is due to using a web crawler for automatic mapping of events.

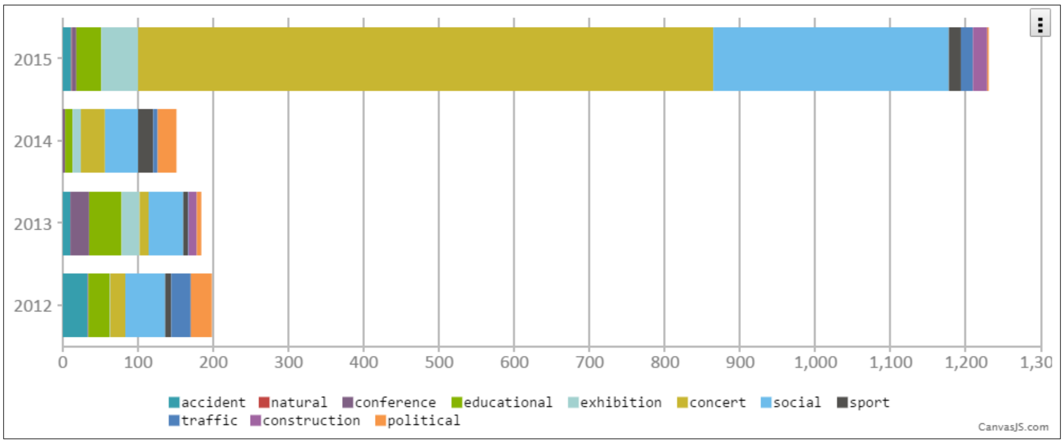

Figure 5-23: The number of mapped events for Munich from 2012 to 2015.

As can be seen in Figure 5-24, the number events in Munich are explored based on category. Obviously, concerts and social events have the highest share among all mapped events by volunteers and the web crawler.

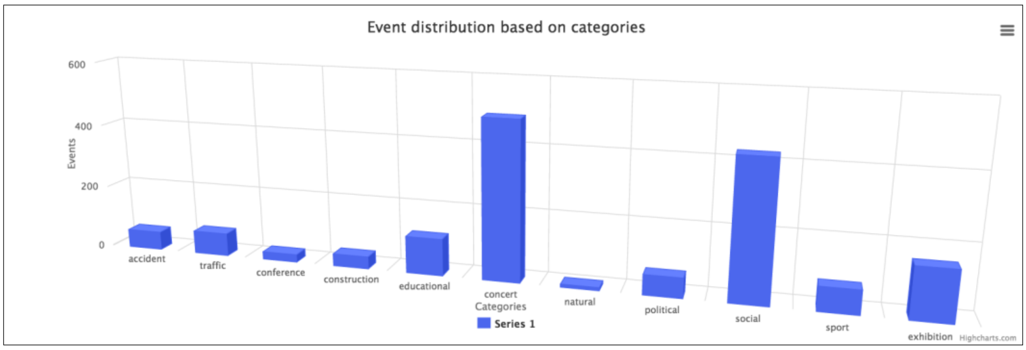

Figure 5-24: The number of mapped events for Munich from 2012 to 2015.

The popularity of events can be analyzed based on their number of participants. Figure 5-25 demonstrates the distribution of all available event based on the number of participants in each season from winter 2014 to winter 2015 in Munich.

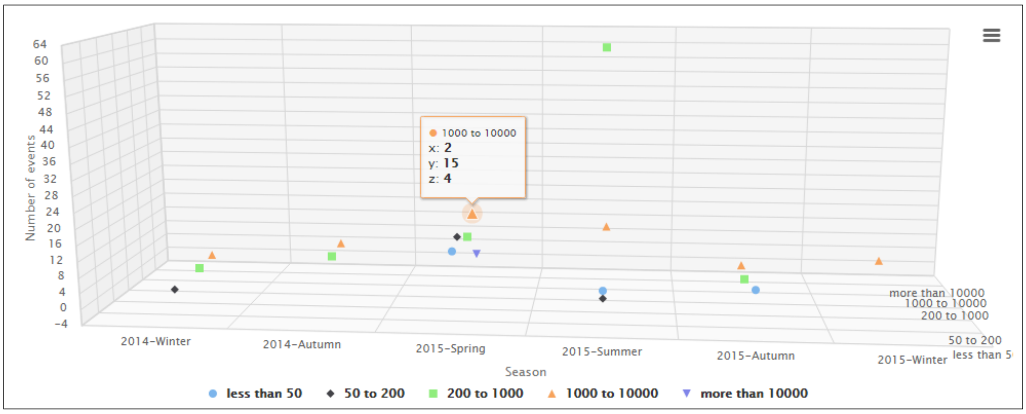

Figure 5-25: The number of participants for events in each season.

Figure 5-26 shows the distribution of events based on number of participants in Munich from 2012 to 2015 in different season. Indeed this illustrates public preferences' for taking part in different seasons.

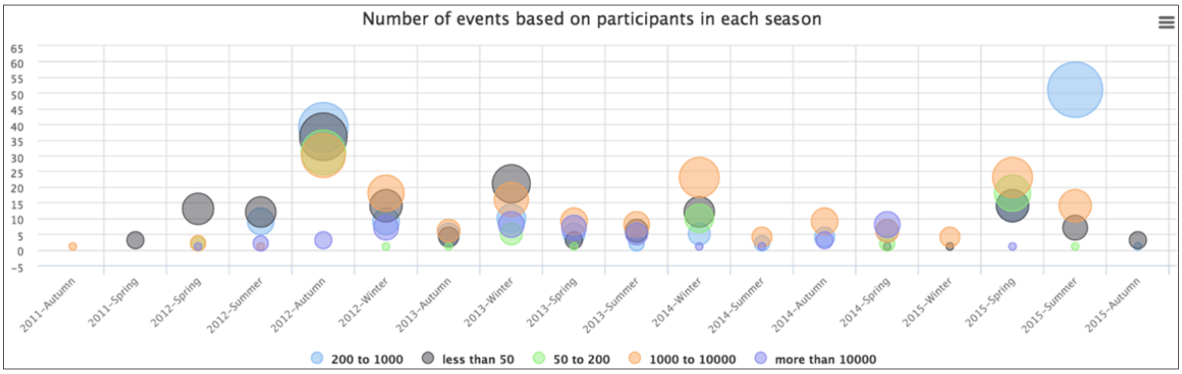

Figure 5-26: The number of mapped events with different number of participants for Munich from 2012 to 2015.

The heat map provides possibility of relationship analysis between two variables. It is a matrix-like presentation of a density map. The color in each cell presents the corresponding value of cross variables. Figure 5-27 shows the heat map (District-Category) of Munich which allows users to explore the relationships among different categories in different districts. In this way, the densest category in each district as well as all districts can be investigated. According to the result (based on the OSM-Event database), the majority of accidents and traffics happen in Neuhausen and Obergiesing while the distribution of other events in different districts is nearly homogeneous

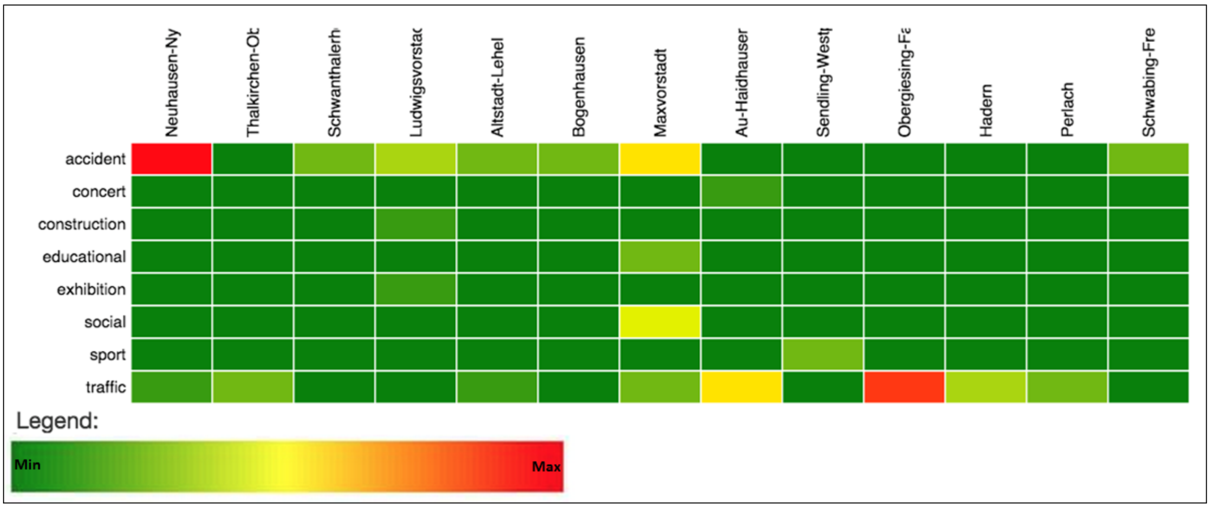

Figure 5-27: Heat map of Munich districts and event categories in February 2013.

6 CONCLUSION AND FURTHER WORK

In this thesis, the concept of "mapping" goes beyond the principle of mapping an object as a conceptual geographic entity, which has a distinct spatial, temporal and attributive identity. The main goal of this thesis is to present a conceptual model for managing geo-knowledge. This can handle real world dynamisms such as the relationships amongst objects with each other, objects with events, events with events and the involving processes. This study uses a generic event-oriented perspective to implicitly represent causal relationships among different components of a Spatio-Temporal Geographic Information System. From this new perspective, the objects in space and time are considered merely as information elements of the events, which are connected to other event elements through internal or external processes. Indeed in this model, the essential components for modeling dynamic phenomena are objects, states, processes and events. Objects belong to states, and states and processes belong to the 'dynamic snapshot' view of the world at one time. As our snapshots contain both processes and objects, they are no longer static but have an inherent dynamism which provides the foundation for generating events. By looking at the snapshots we can see different objects, in various states, which undergoing particular processes. These snapshots are constantly renewed as time passes; the snapshots alter from one moment to the next, because the present elements in the snapshot can undergo change. Whereas, events are fixed historical records, which are not renewed or replaced by a new record, however as time passes and events happen, they are gradually added to the record. The approach can be used to model events at any scale; the core or any of the sub events of our example can be for instance the event of interest. Indeed any event that is represented in the model is a sequence of sub-events into which different processes are resolved when checked at a finer granularity level. In this study, the relationship between objects and events with *1* to *N*, *N* to *1* and *N* to *M* was mapped in space and time and the internal and external relationship amongst event with other events and processes was not considered in implementation. An interesting field of research in this area is to develop a sophisticated ontology for representing the interaction and causal relationship between occurring events and involving external and internal processes of each type of event that might happen in our cities at different scales. In addition, the aggregation of the process and event and the corresponding semantic information is a vital task. This ontology can be used to systematically define events, processes and their relationships. In this way, the dynamic face of the globe along with its historical record can be modeled and the map will be a knowledge-based system to understand the interaction of its component, to predict the future and design the required services and facilities.

Due to unique opportunity offered by the concept of VGI, a framework was developed to provide the possibility of collecting and storing event-related information elements in OSM-platform. In order to distinguish and identify VGI, it is essential to specify first a set of parameters and criteria that allows evaluation of the similarities and differences across the VGI panel, and ultimately determine the major sorts of VGI. These parameters mainly originate from existing models and frameworks in the PPGIS, neogeography as well as from the critical analysis of the VGI landscape. In the conceptual VGI framework, the parameters are organized in four categories: enablers, context, mechanisms, and utilization and consume. Each category has in its turn various parameters to characterize and identify different VGI types. Enablers – unlike other categories – are not a set of parameters or criteria for describing different characteristics of VGI. VGI inherits indeed many of its unique features from scientific and technological enablers such as web 2.0, GIS and PPGIS that have been extensively described in chapter two. To fully understand VGI, we should admit and understand the relationship of enablers to their phenomenal offspring VGI. Hence, the foundation of the proposed framework has been shown as scientific and technological enablers. Context describes the essential distinctions in the nature of the various types of VGI. This category represents various VGI continuums. Mechanisms represent different approaches and tools that facilitate contributions. And finally utilization describes the last phase of the VGI cycle - data utilization and synthesis that leads to decision based on the VGI information.

In the implementation phase, the OSM database and technology infrastructure was extended to explore the idea of temporal data collection and storage in the same OSM format. Event related information are collected and entered into the database using special tags. The model provides a standard way to mathematically model the changing world and a firm basis for the logical modelling of dynamical systems due to considering processes in the snapshots. Based on the event model four mechanisms were developed to facilitate the procedure of event data collection either by contributors or automatically. These four mechanisms are as follow; a) a plugin for JOSM offline editor, b) an Android application, c) a web crawler for automatic collection, and d) a social media monitoring framework. The developed event definition can capture spatio-temporal and semantic information of object and Event while answering to the proposed questions: a)“What” for all desired semantic information like type or class of events b)“Where” for location of events by assigning involving objects (temporal information as well as semantic information explicitly are collected through object mapping in OSM platform), c)“When” for time point or duration of events, d) “How” for event frequency or pattern, f) “Who” organization, agent, external force. This collected information is stored in a local database (so called Event-OSM database). The database is compatible with the original OSM database and is frequently synchronized with OSM server. OSM has predefined set of tags to describe spatial data. To fully facilitate the mapping of the N to M relationships between objects and events, two extra strategies were considered in developing the syntax. The first strategy is; namespacing in tags and assigning an index to each event. This enables contributors to assign any different number of events to the same location or object, which is very common for locations like exhibitions or sport arenas. Second, we facilitate the assignment of different spatial objects to a same event. This will avoid redundancy in our data base, which is a baseline for any record from all event collection tools. Besides increasing the software and infrastructure maturity, future work may include incorporating more event tags in the syntax, in order to fully model all aspects of events such as processes and involving relationships in the OSM database. Another interesting investigation can be developing a proper data model and database system to better represent spatio-temporal information for events. Further step is to develop sophisticated analytical visualization tools for events which enable us to analyze dynamism of cities.

To utilize the collected event information, a Volunteered Location Based Service (VLBS) named Openeventmap was developed. The service provides users the possibility of mining the event database and answers the spatial, temporal and attributive queries associated with the proposed event definition. As an extension to “openeventmap.de” and in order to further enrich the collected event data base applications and tools can be developed to search in different social media and collect information such as tweets and pictures related to the desired event. Furthermore, the geo-tagged pictures and tweets can be mapped on OSM to see where the social media information comes from, and what is the reaction of public on any kind of events mapped in OSM data base or social media streams. Another interesting extension of the work could be to automate the whole collection processes in the developed framework which requires data fusion techniques, semantic analysis, and automatic or semi-automatic quality check for detected events.

The last part of this research is devoted to changing the role of the volunteers from pure sensors in the VGI concept. In this perspective, volunteers not only observe environment and collect the information through their senses, but also are able to understand the meaning of the data they gathered. For this purpose an Event Visualization and Analysis Tool (EVT) was developed. This tool was designed to increase the general awareness of the volunteers in spatial analysis and shift their power and influences from a mere data collector to a knowledge producer. This accelerates the transition of mastery from professional cartographers to the public; indeed everybody who can understand the space will be able to collect, analyze, evolve and finally synthesize the spatial knowledge. An important extension for this work is to investigate the feedback of volunteers in using the tool. This will help to further improve the tool and increase the user friendliness of features. Furthermore, exploring the patterns in using the tool may increase our understanding of user's preferences about space. While developing technological infrastructures of the platform is necessary to vitalize the city maps, applying strategies to maximize the engagement of volunteers is of vital importance. These strategies require a comprehensive understanding of various factors such as volunteer's motives, cultural background, their expectations and interests. In addition, managing highly diverse volunteer community and facilitating a productive relationship among community members is another big challenge that can affect achievement of the goal.

REFERENCES

Abraham, T., Roddick, J.F., 1999. Survey of spatio-temporal databases. Geoinformatica 3, 61–99.

Agarwal, P., 2005. Ontological Considerations in GIScience. Int J Geogr Inf Sci 19, 501–536.

Albors, J., Ramos, J.C., Hervas, J.L., 2008. New learning network paradigms: Communities of objectives, crowdsourcing, wikis and open source. Int J Inf Manage 28, 194–202.

Allen, E., Edwards, G., Bédard, Y., 1995. Qualitative causal modeling in temporal GIS. In: Spatial Information Theory. pp. 397–412.

Allen, J.F., 1984. Towards a general theory of action and time. Artif Intell 23, 123–154.

Allen, N., Ingham, J., Johnson, B., Merante, J., Noveck, B.S., Stock, W., Tham, Y., Webbink, M., Wong, C., 2008. First anniversary report, New York. New York, New York, USA.

Ambati, V., Vogel, S., Carbonell, J., 2010. Active Learning and Crowd-Sourcing for Machine Translation. In: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (Eds.), Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). European Language Resources Association (ELRA), pp. 2169–2174.

Andriole, S., 2010. Business Impact of Web 2.0 Technologies. Commun ACM 53, 67–79.

Antikainen, M.J., Vaataja, H.K., 2010. Rewarding in open innovation communities – how to motivate members. Int J Entrep Innov Manag 11, 440–456.

Archak, N., 2010. Money, glory and cheap talk: analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on TopCoder.com. In: Proceedings of the 19th International Conference on World Wide Web - WWW '10. ACM, New York, NY, USA, p. 21.

Armstrong, M.P., 1988. Temporality in Spatial Databases. In: Accessing the World, Vol. II, Falls Church, VA, USA: American Society for Photogrammetry and Remote Sensing. pp. 880–889.

Arnstein, S.R., 1969. A Ladder Of Citizen Participation. J Am Inst Plann 35, 216–224.

Baker, A.R.H., 2003. Geography and history: Bridging the divide, Cambridge Studies in Historical Geography. Cambridge University Press, Cambridge, England.

Balashov, Y., 1999. Relativistic Objects. Noûs 33, 644–662.

Balram, S., 2006. Collaborative Geographic Information Systems. Idea Group Inc, London, United Kingdom.

Barrington, L., O'Malley, D., Turnbull, D., Lanckriet, G., 2009. User-centered design of a social game to tag music. In: Proceedings of the ACM SIGKDD Workshop on Human Computation HCOMP 09. ACM Press, New York, NY, USA, pp. 7–10.

Beard, K., 2006. Modelling Change in Space and Time: An Event-Based Approach. In: Drummond, J., Billen, R., João, E., Forrest, D. (Eds.), Dynamic and Mobile GIS: Investigating Changes in Space and Time. CRC Press, Boca Raton, FL, pp. 55–76.

Bennett, J., 2002. What Events Are. In: Casati, R., Varzi, A.C. (Eds.), Events. Dartmouth, Aldershot, pp. 137–151.

Bichler, M., Gupta, A., Ketter, W., 2010. Designing smart markets. Inf Syst Res 21, 688–699.

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

Black, N., 2007. OpenStreetMap - Geodata Collection for the 21st Century. In: AGI Annual

Conference.

Bojórquez-Tapia, L.A., Diaz-Mondragón, S., Ezcurra, E., 2001. GIS-based approach for participatory decision making and land suitability assessment. Int J Geogr Inf Sci.

Bonabeau, E., 2001. Swarm Intelligence, A Whole New Way to Think About Business. Harv Bus Rev 79, 106–114.

Bonabeau, E., 2009. Decisions 2.0 : The Power of Collective Intelligence. MIT Sloan Manag Rev 50, 45–52.

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., Shirk, J., 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. Bioscience 59, 977–984.

Boudreau, K.J., Lacetera, N., Lakhani, K.R., 2011. Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. Manage Sci 57, 843–863.

Brabham, D.C., 2008a. Mooving the crowed at iStockphoto : The composition of the crowd and motivations for participation in a crowdsourcing application. First Monday 13, 1–19.

Brabham, D.C., 2008b. Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. Converg Int J Res into New Media Technol 14, 75–90.

Brabham, D.C., 2010. MOVING THE CROWD AT THREADLESS. Information, Commun Soc.

Bretschneider, U., Rajagopalan, B., Leimeister, J.M., 2011. Idea generation in Virtual Communities for Innovation: The influence of participants' motivation on idea quality. In: Proceedings of the Annual Hawaii International Conference on System Sciences (HICSS). CPS, Grand Wailea, Maui, pp. 3467–3479.

Budhathoki, N.R., 2010. Participants' Motivations to Contribute Geographic Information in an Online Community. University of Illinois at Urbana-Champaign.

Budhathoki, N.R., Bruce, B., Nedovic-Budic, Z., 2008. Reconceptualizing the role of the user of spatial data infrastructure. GeoJournal.

Budhathoki, N.R., Nedović-Budić, Z., Bruce, B., 2010. An interdisciplinary frame for understanding volunteered geographic information. Geomatica.

Bushmann, F., Meunier, R., Rohnert, H., 1996. Pattern-oriented software architecture: A system of patterns. John Wiley&Sons 1, 476.

Carver, S., 2003. The future of participatory approaches using geographic information: Developing a research agenda for the 21st century. URISA J 15, 61–71.

Casati, R., 2005. Commonsense, Philosophical and Theoretical Notions of an Object: Some Methodological Problems. Monist 88, 571–599.

Casati, R., Varzi, A.C., 2008. Event concepts. In: Shipley, T.F., Zacks, J. (Eds.), Understanding Events: From Perception to Action. Oxford University Press, New York, NY, USA, pp. 31–54.

Chanal, V., Caron-Fasan, M.-L., 2008. How to invent a new business model based on crowdsourcing : the Crowdspirit ® case. In: Conférence de l'Association Internationale de Management Stratégique. HAL-SHS, Sophia-Antipolis, France, pp. 1–27.

Chen, J., Jiang, J., 2000. An Event-Based Approach to Spatio-Temporal Data Modeling in Land Subdivision Systems. Geoinformatica 4, 387–402.

Chesbrough, H.W., 2003. Open Innovation: The New Imperative for Creating and Profiting from Technology. Harvard Business Review Press, Boston, MA, USA.

Chesbrough, H.W., 2006. Open Business Models: How to Thrive in the New Innovation Landscape.

Harvard Business School Press.

Chilton, S., 2009. Crowdsourcing Is Radically Changing the Geodata Landscape : Case Study of Openstreetmap. In: 24th International Cartographic Conference, Chile. pp. 1–7.

Chrisman, N.R., 1998. Beyond the snapshot: changing the approach to change, error, and process. In: Egenhofer, M.J., Golledge, R.G. (Eds.), Spatial and Temporal Reasoning in Geographic Information Systems. Oxford University Press, New York, NY, USA, pp. 85–93.

Chrisman, N.R., McGranaghan, M., 1990. Accuracy of Spatial Databases. In: Goodchild, M.F., Kemp, K.K. (Eds.), NCGIA Core Curriculum in GIS. National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA, Santa Barbara, CA, USA.

Claramunt, C., Thériault, M., 1995. Managing Time in GIS An Event-Oriented Approach. In: Recent Advances in Temporal Databases. pp. 23–42.

Clark, M.L., Aide, T.M., 2011. Virtual interpretation of Earth Web-Interface Tool (VIEW-IT) for collecting land-use/land-cover reference data. Remote Sens 3, 601–620.

Clary, E.G., Snyder, M., Ridge, R.D., Copeland, J., Stukas, A.A., Haugen, J., Miene, P., 1998. Understanding and assessing the motivations of volunteers: a functional approach. J Pers Soc Psychol 74, 1516–1530.

Clinton, B., 2000. Improving the Civilian Global Positioning System (GPS) [WWW Document]. The White House. URL http://clinton4.nara.gov/textonly/WH/EOP/OSTP/html/0053_4.html (accessed 6.1.15).

Coleman, D.J., Georgiadou, Y., Labonte, J., Observation, E., Canada, N.R., 2009. Volunteered Geographic Information : The Nature and Motivation of Produsers. Int J Spat Data Infrastructures Res 4, 332–358.

Cooper, A.K., Coetzee, S., Kaczmarek, I., Kourie, D.G., Iwaniak, A., 2011. Challenges for quality in volunteered geographical information. In: AfricaGEO 2011. Cape Town, South Africa, p. 13.

Coote, A., Rackham, L., 2008. Neogeography Data Quality - is it an issue? In: Proceedings of AGI Geocommunity ’08. Stratford-Upon-Avon, UK, p. 1.

Craig, W.J., Harris, T.M., Weiner, D., 2002. Community participation and geographic information systems, Community participation and geographic information systems. Taylor & Francis Inc, New York.

Crampton, J.W., 2003. Cartographic Rationality and the Politics of Geosurveillance and Security. Cartogr Geogr Inf Sci 30, 135–148.

Crone, G.R., 1968. Maps and their makers: an introduction to the history of cartography by G.R. Crone., Fourth. ed. Hutchinson & Co LTD, London, UK.

Crutcher, M., Zook, M., 2009. Placemarks and waterlines: Racialized cyberscapes in post-Katrina Google Earth. Geoforum 40, 523–534.

Dana, P.H., 2008. Surveys of People and Place. In: Wilson, J.P., Fotheringham, A.S. (Eds.), The Handbook of Geographic Information Science. Wiley, pp. 494–518.

Davidson, D., 1967. Causal Relations. J Philos 64, 691.

Deci, E.L., Ryan, R.M., 1985. Intrinsic motivation and self-determination in human behavior, Perspectives in Social Psychology. Springer Verlag US.

Deparday, V., 2010. Enhancing Volunteered Geographical Information (VGI) Visualization with Open Source Web-Based Software by. Univ Waterloo.

DeSanctis, G., Poole, M.S., 1994. Capturing the Complexity in Advanced Technology Use: Adaptive

Structuration Theory. Organ Sci 5, 121–147.

Dholakia, U.M., Bagozzi, R.P., Pearo, L.K., 2004. A social influence model of consumer participation in network- and small-group-based virtual communities. Int J Res Mark 21, 241–263.

Di Gangi, P.M., Wasko, M., 2009. Steal my idea! Organizational adoption of user innovations from a user innovation community: A case study of Dell IdeaStorm. Decis Support Syst 48, 303–312.

Dickinson, J.L., Zuckerberg, B., Bonter, D.N., 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. Annu Rev Ecol Evol Syst 41, 149–172.

Doan, A., Ramakrishnan, R., Halevy, A.Y., 2011. Crowdsourcing systems on the World-Wide Web. Commun ACM 54, 86–96.

Dobson, J.E., Fisher, P.F., 2006. Panopticon III: Framing the National Debate over Human Tracking. In: American Association for the Advancement of Science. in the Orwell’s Wolf Is Back at the Door: Tracking Kids, Dogs, Old People, and Everybody in Between. St. Louis, MO.

Dodgson, M., Gann, D., Salter, A., 2006. The role of technology in the shift towards open innovation: The case of Procter & Gamble. R D Manag 36, 333–346.

Dunn, C.E., 2007. Participatory GIS : a people’s GIS ? Prog Hum Geogr 31, 616–637.

Ebner, W., Leimeister, J.M., Krcmar, H., 2009. Community engineering for innovations: The ideas competition as a method to nurture a virtual community for innovations. R D Manag 39, 342–356.

Eisnor, D.-A., 2006. What is neogeography anyway? [WWW Document]. Platial News Neogeography. URL http://platial.typepad.com/news/2006/05/what_is_neogeog.html

Elwood, S., 2002. The impacts of GIS use for neighborhood revitalization in Minneapolis. In: Harris, T.M.., Craig, W.J.., Weiner, D. (Eds.), Community Participation and Geographic Information Systems. CRC Press, pp. 77–88.

Elwood, S., 2008a. Volunteered geographic information: Key questions, concepts and methods to guide emerging research and practice. GeoJournal.

Elwood, S., 2008b. Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS. GeoJournal 72, 173–183.

Elwood, S., 2008c. Geographic Information Science: new geovisualization technologies -- emerging questions and linkages with GIScience research. Prog Hum Geogr 33, 256–263.

Enkel, E., Gassmann, O., Chesbrough, H., 2009. Open R&D and open innovation: Exploring the phenomenon. R D Manag 39, 311–316.

Erickson, L.B., 2013. Hanging with the right crowd: Crowdsourcing as a new business practice for innovation, productivity, knowledge capture, and marketing. ProQuest Diss Theses.

Ester, M., Kriegel, H.P., Sander, J., Xu, X., 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: Second International Conference on Knowledge Discovery and Data Mining. pp. 226–231.

Evans, A.J., Kingston, R., Carver, S., 2004. Democratic input into the nuclear waste disposal problem: The influence of geographical data on decision making examined through a Web-based GIS. J Geogr Syst 6, 117–132.

Faymonville, P., Wang, K., Miller, J., Belongie, S., 2009. CAPTCHA-based image labeling on the Soylent Grid. In: Proceedings of the ACM SIGKDD Workshop on Human Computation HCOMP 09. ACM, New York, NY, USA, pp. 46–49.

Feick, R., Roche, S., 2013. Understanding the value of VGI. In: Sui, D., Elwood, S., Goodchild, M. (Eds.), Volunteered Geographic Information (VGI) in Theory and Practice. Springer, Dordrecht Heidelberg New York London, pp. 15–29.

Feller, J., Finnegan, P., Hayes, J., O'Reilly, P., 2009. Institutionalising information asymmetry: governance structures for open innovation. Inf Technol People 22, 297–316.

Felstinerf, A., 2011. Working the Crowd : Employment and Labor Law in the Crowdsourcing Industry. Berkeley J Employ Labor Law 32, 143–204.

Fine, K., 2008. In Defence of Three-Dimensionalism. R Inst Philos Suppl 62, 1–16.

Flanagin, A.J., Metzger, M.J., 2007. The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. New Media Soc 9, 319–342.

Flanagin, A.J., Metzger, M.J., 2008. The credibility of volunteered geographic information. GeoJournal.

Floyd, I.R., Cameron Jones, M., Rathi, D., Twidale, M.B., 2007. Web mash-ups and patchwork prototyping: user-driven technological innovation with web 2.0 and open source software. In: Proceedings of the Annual Hawaii International Conference on System Sciences. Citeseer, Waikoloa, Big Island, Hawaii, p. 1454.

Foundation, Jq., 2014. jQuery [WWW Document]. URL http://www.jqueryui.com (accessed 10.30.14).

Fredberg, T., Piller, F.T., 2011. The paradox of tie strength in customer relationships for innovation: A longitudinal case study in the sports industry. R D Manag 41, 470–484.

Füller, J., 2010. Refining Virtual Co-Creation from a Consumer Perspective. Calif Manage Rev.

Füller, J., Hutter, K., Faullant, R., 2011. Why co-creation experience matters? Creative experience and its impact on the quantity and quality of creative contributions. R D Manag 41, 259–273.

Galton, A., 2000. Qualitative Spatial Change. Oxford University Press, Oxford.

Galton, A., 2004. Fields and Objects in Space, Time, and Space-time. Spat Cogn Comput 4, 39–68.

Galton, A., 2005. Ontology is not just about objects: Poorly individuated phenomena in geo-ontology. Appl Ontol 1, 47–52.

Galton, A., 2006a. On what goes on: The ontology of processes and events. In: Bennett, B., Fellbaum, C. (Eds.), Formal Ontology in Information Systems. IOS Press, Amsterdam, Netherlands, pp. 4–11.

Galton, A., 2006b. Processes as Continuants. In: Thirteenth International Symposium on Temporal Representation and Reasoning (TIME'06). IEEE, Budapest, pp. 187–187.

Galton, A., 2012. States, Processes and Events, and the Ontology of Causal Relations. In: Donnelly, M., Guizzardi, G. (Eds.), Formal Ontology in Information Systems. IOS Press. IOS Press, Amsterdam, Netherlands, pp. 279–292.

Galton, A., Mizoguchi, R., 2009. The water falls but the waterfall does not fall : New perspectives on objects , processes and events. Appl Ontol 4, 71–107.

Galton, A., Worboys, M., 2005. Processes and events in dynamic geo-networks. In: Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 45–59.

Ghose, R., 2001. Use of Information Technology for Community Empowerment: Transforming Geographic Information Systems into Community Information Systems. Trans GIS 5, 141–163.

Gibson, R., Erle, S., 2006. Google Maps Hacks. O'Reilly Media, Inc.

Girres, J.F., Touya, G., 2010. Quality Assessment of the French OpenStreetMap Dataset. Trans GIS 14, 435–459.

Goodchild, M., 2008. Assertion and authority: the science of user-generated geographic content. In: Proceedings of the Colloquium for Andrew U. Frank's 60th Birthday . GeoInfo 39. Department of Geoinformation and Cartography, Vienna University of Technology, Vienna. Vienna, Austria.

Goodchild, M.F., 2007a. Citizens as sensors: The world of volunteered geography. GeoJournal 69, 211–221.

Goodchild, M.F., 2007b. Citizens as voluntary sensors: Spatial data infrastructure in the world of Web 2.0. Int J Spat Data Infrastructures Res 2, 24–32.

Goodchild, M.F., 2007c. CITIZENS AS SENSORS: WEB 2.0 AND THE VOLUNTEERING OF GEOGRAPHIC INFORMATION. GeoFocus 7, 8–10.

Goodchild, M.F., 2008a. Commentary: Whither VGI? GeoJournal.

Goodchild, M.F., 2008b. Spatial accuracy 2.0. In: Proceeding of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Shanghai, China, pp. 1–7.

Goodchild, M.F., 2010. Twenty years of progress: GIScience in 2010. J Spat Inf Sci 1, 3–20.

Goodchild, M.F., Glennon, J.A., 2010. Crowdsourcing geographic information for disaster response: a research frontier. Int J Digit Earth 3, 231–241.

Goodchild, M.F., Li, L., 2012. Assuring the quality of volunteered geographic information. Spat Stat 1, 110–120.

Grenier, L., 1998. Working with indigenous knowledge: a guide for researchers, International Development Research Centre. Ottawa, Cairo, Dakar, Johannesburg, Montevideo, Nairobi, New Delhi, Singapore.

Grenon, P., Smith, B., 2004. SNAP and SPAN: Towards dynamic spatial ontology. Spat Cogn Comput 1, 69–103.

Grossner, K., 2010. Event Objects for Spatial History. In: Purves, R., Weibel, R. (Eds.), Extended Abstracts Volume, GIScience 2010. Zurich.

Hägerstrand, T., 1970. What about people in Regional Science? Pap Reg Sci Assoc 24, 6–21.

Haklay, M., 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. Environ Plan B Plan Des 37, 682–703.

Haklay, M., Antoniou, V., Basiouka, S., Mooney, P., 2014. Crowdsourced Geographic Information Use in Government. London, United Kingdom.

Haklay, M., Basiouka, S., Antoniou, V., Ather, A., 2010. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. Cartogr J 47, 315–322.

Haklay, M., Singleton, A., Parker, C., 2008. Web mapping 2.0: The neogeography of the GeoWeb. Geogr Compass 2, 2011–2039.

Haklay, M., Weber, P., 2008. OpenStreet map: User-generated street maps. IEEE Pervasive Comput 7, 12–18.

Halavais, A., 2009. DO DUGG DIGGERS DIGG DILIGENTLY? Information, Commun Soc 12, 444–459.

Hall, B.H., Leahy, M.G., 2011. Map-chatting within the geospatial web. In: Veenendaal, B. (Ed.), Advances in Web-Based GIS, Mapping Services and Applications. CRC Press, pp. 229–254.

Hall, G.B., Chipeniuk, R., Feick, R.D., Leahy, M.G., Deparday, V., 2010. Community-based production of geographic information using open source software and Web 2.0. Int J Geogr Inf Sci.

Hall, G.B., Moore, A., Knight, P., Hankey, N., 2009. The extraction and utilization of local and scientific geospatial knowledge within the Bluff oyster fishery, New Zealand. J Environ Manage 90, 2055–2070.

Hardey, M., 2007. The city in the age of web 2.0 a new synergistic relationship between place and people. Information, Commun Soc.

Hars, A., Ou, S., 2002. Working for Free? Motivations for Participating in Open-Source Projects. Int J Electron Commer 6, 25–39.

Harvey, F., 2007. Just another private-public partnership? Possible constraints on scientific information in virtual map browsers. Environ Plan B Plan Des 34, 761–764.

Haythornthwaite, C., 2009. Crowds and communities: Light and heavyweight models of peer production. In: Proceedings of the 42nd Annual Hawaii International Conference on System Sciences, HICSS. Big Island, HI, pp. 1–10.

Haythornthwaite, C., 2011. Learning Networks, Crowds and Communities. In: Learning Analytics and Knowledge. pp. 18–22.

Heer, J., Bostock, M., 2010. Crowdsourcing graphical perception. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI ’10. Atlanta, GA; USA, pp. 203–212.

Heller, M., 2008. The ontology of physical objects: four-dimensional hunks of matter, Reissue ed. ed. Cambridge University Press.

Hertel, G., Niedner, S., Herrmann, S., 2003. Motivation of Software Developers in Open Source Projects: An Internet Based Survey of Contributors to the Linux Kernel. Res Policy 32, 1159–1177.

Hess, C., Ostrom, E., 2007. A Framework for Analyzing the Knowledge Commons. In: Hess, C., Ostrom, E. (Eds.), Understanding Knowledge as a Commons: From Theory to Practice. MIT Press, Cambridge, Massachusetts, London, England, pp. 41–81.

Ho, C., Chang, T., Lee, J., Hsu, J.Y., Chen, K., 2009. KissKissBan: a competitive human computation game for image annotation. In: Proceedings of the ACM SIGKDD Workshop on Human Computation - HCOMP ’09. ACM, New York, NY, USA, pp. 11–14.

Hornsby, K., Egenhofer, M.J., 2000. Identity-based change: a foundation for spatio-temporal knowledge representation. Int J Geogr Inf Sci 14, 207–224.

Hornsby, K.S., Cole, S., 2007. Modeling moving geospatial objects from an event-based perspective. Trans GIS 11, 555–573.

Houle, B.J., Sagarin, B.J., Kaplan, M.F., 2005. A Functional Approach to Volunteerism: Do Volunteer Motives Predict Task Preference? Basic Appl Soc Psych 27, 337–344.

Howe, J., 2006. Crowdsourcing: A Definition [WWW Document]. Crowdsourcing Track rise Amat. URL http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html (accessed 5.1.15).

Howe, J., 2008. Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business, 1st ed. Crown Publishing Group, New York, NY, USA.

Hristova, D., Quattrone, G., Mashhadi, A., Capra, L., 2013. The Life of the Party: Impact of Social Mapping in OpenStreetMap. In: Seventh International AAAI Conference on Weblogs and Social Media. Cambridge, Massachusetts, USA.

Huberman, B.A., Subrahmanyam, A., Romero, D.M., Wu, F., 2009. Crowdsourcing, attention and productivity. J Inf Sci 35, 758–765.

Hudson-Smith, A., Batty, M., Crooks, A., Milton, R., 2009. Mapping for the Masses Accessing Web 2.0 Through Crowdsourcing. Soc Sci Comput Rev 27.

Jana, R., 2009. How Intuit Makes a Social Network Pay - BusinessWeek [WWW Document]. Bloom Businessweek Mag. URL http://www.businessweek.com/magazine/content/09_28/b4139066365300.htm (accessed 5.1.15).

Jankowski, P., Nyerges, T., 2001. GIS-Supported Collaborative Decision Making: Results of an Experiment. Ann Assoc Am Geogr 91, 48–70.

Jeppesen, L.B., Frederiksen, L., 2006. Why Do Users Contribute to Firm-Hosted User Communities? The Case of Computer-Controlled Music Instruments. Organ Sci 17, 45–63.

Jeppesen, L.B., Lakhani, K.R., 2010. Marginality and Problem-Solving Effectiveness in Broadcast Search. Organ Sci 21, 1016–1033.

Johnson, W.H.A., Filippini, R., 2009. Internal vs. external collaboration: What works. Res Technol Manag 52, 15–17.

Jouret, G., 2009. Inside cisco's search for the next big idea. Harv Bus Rev 87, 43.

Kittur, A., Chi, E., Pendleton, B.A., Suh, B., Mytkowicz, T., 2007. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. In: Algorithmica. San Jose, CA.

Klippel, A., Worboys, M., Duckham, M., 2006. Geographic event conceptualization. Cogn Process 7.

Klippel, A., Worboys, M., Duckham, M., 2008. Identifying factors of geographic event conceptualisation. Int J Geogr Inf Sci 22, 183–204.

Kollock, P., 1999. The economies of online cooperation: Gifts and public goods in cyberspace. (1 of 2). In: Smith, M., Kollock, P. (Eds.), Communities in Cyberspace. Routledge, London, UK, pp. 220–242.

Kozinets, R. V., Hemetsberger, A., Schau, H.J., 2008. The Wisdom of Consumer Crowds: Collective Innovation in the Age of Networked Marketing. J Macromarketing.

Kuznetsov, S., 2006. Motivations of contributors to Wikipedia. ACM SIGCAS Comput Soc.

Kyem, P.A.K., 2004. Of intractable conflicts and participatory GIS applications: The search for consensus amidst competing claims and institutional demands. Ann Assoc Am Geogr.

Labbe, E., 2002. A Conceptual Object-Oriented Model for Complex Geographical Entities – the DDCO model –. Vrije Universiteit Brussel - Belgium.

Laituri, M., Kodrich, K., 2008. On line disaster response community: People as sensors of high magnitude disasters using internet GIS. Sensors 8, 3037–3055.

Lake, R.W., 1993. Planning and applied geography: positivism, ethics, and geographic information systems. Prog Hum Geogr 17, 404–413.

Lakhani, K., Wolf, R.G., 2003. Why hackers do what they do: Understanding motivation and effort in free/open source software projects, MIT Sloan Working Paper No. 4425-03.

Lakhani, K.R., Lohse, P. a, Panetta, J. a, Jeppesen, L.B., 2007. The Value of Openness in Scientific Problem Solving, Harvard Business School Working Paper. Boston, MA, USA.

Lakhani, K.R., Panetta, J.A., 2007. The Principles of Distributed Innovation. Innov Technol Governance, Glob 2, 97–112.

Langran, G., 1992. Time in Geographic Information Systems. Taylor & Francis Inc, London.

Langran, G., Chrisman, N.R., 1988. A Framework For Temporal Geographic Information. Cartogr Int J Geogr Inf Geovisualization 25, 1–14.

Leimeister, J.M., Huber, M., Bretschneider, U., Krcmar, H., 2009. Leveraging Crowdsourcing: Activation-Supporting Components for IT-Based Ideas Competition. J Manag Inf Syst 26, 197–224.

Lewis, D., 1986. Events. Philos Pap II, 241–269.

Liu, Y., Mcgrath, R.E., Wang, S., Pietrowicz, M., Futrelle, J., Myers, J.D., 2008. Towards A Spatiotemporal Event-Oriented Ontology. In: Artificial Intelligence. Indianapolis, IN, USA, pp. 18–20.

Lo, C.P., Yeung, A.K.W., 2007. Concepts and techniques of geographic information systems. Pearson Prentice Hall, Upper Saddle River, NJ.

Maguire, D.J., 2007. GeoWeb 2.0 and volunteered GI. Work Volunt Geogr Information, St Barbar CA.

Mahemoff, M., 2006. Ajax Design Patterns, Designing Information Systems Festschrift für Jürgen Krause. O'Reilly Media, Inc.

Malone, T.W., Laubacher, R., Dellarocas, C., 2009. Harnessing crowds : Mapping the genome of collective intelligence. MIT Sloan Res Pap 4732-09, 1–20.

Mark, D., 1998. Geospatial lifelines, Integrating Spatial and Temporal Databases, Dagstuhl Seminars.

Maslow, A.H., 1987. Motivation and Personality, 3rd ed. Longman.

Mau, I., Hornsby, K.S., Bishop, I.D., 2007. Modeling Geospatial Events and Impacts Through Qualitative Change. In: Barkowsky, T., Knauff, M., Ligozat, G., Montello, D.R. (Eds.), Spatial Cognition V: Reasoning, Action, Interaction. Springer, Springer Berlin Heidelberg, pp. 156–174.

McAfee, A., 2009. Enterprise 2.0: new collaborative tools for your organization's toughest challenges, 1st ed, Harvard Business Press. Harvard Business Review Press.

Metzger, M.J., Flanagin, A.J., 2011. Using web 2.0 technologies to enhance evidence-based medical information. J Health Commun 16 Suppl 1, 45–58.

Mill, J.S., 1843. A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation. harrison and co. printers, London, UK.

Miller, C.C., 2006. A Beast in the Field: The Google Maps Mashup as GIS/2. Cartogr Int J Geogr Inf Geovisualization.

Miller, H.J., 2003. What about people in geographic information science? Comput Environ Urban Syst 27, 447–453.

Moens, M., Steedman, M., 1988. Temporal ontology and temporal reference. Comput Linguist 14, 15–28.

Muhdi, L., Boutellier, R., 2011. Motivational factors affecting participation and members in two different Swiss innovation communities. Int J Innov Manag 15, 543–562.

Mummidi, L.N., Krumm, J., 2008. Discovering points of interest from users' map annotations. GeoJournal.

Murugesan, S., 2007. Understanding Web 2.0. IT Prof 9, 34–41.

Myers, D.G., 2002. Intuition: Its Powers and Perils. Yale University Press, New Haven, Connecticut.

Newman, G., Zimmerman, D., Crall, A., Laituri, M., Graham, J., Stapel, L., 2010. User-friendly web mapping: lessons from a citizen science website. Int J Geogr Inf Sci 24, 1851–1869.

NumPy-Developers, 2014. NumPy [WWW Document]. URL www.numpy.org

Oliphant, T.E., 2007. Python for scientific computing. Comput Sci Eng 9, 10–20.

Olson, M., 1965. The Logic of Collective Action: Public Goods and the Theory of Groups, 20th ed, Contemporary Sociological Theory, Blackwell, Oxford. Harvard University Press, Cambridge. Massachusetts; London, England.

Openshaw, S., 1991. A view on the GIS crisis in geography, or, using GIS to put Humpty-Dumpty back together again. Environ Plan A 23, 621–628.

Ortega, F., Gonzalez Barahona, J.M., 2007. Quantitative analysis of thewikipedia community of users. In: Proceedings of the 2007 International Symposium on Wikis WikiSym 07. pp. 75–86.

OSM Community, 2015. OpenStreetMap Wiki [WWW Document]. OpenStreetMap Wiki. URL http://wiki.openstreetmap.org/ (accessed 6.1.15).

Page, S.E., 2008. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton University Press.

Parameswaran, M., Whinston, A.B., 2007. Social Computing: An Overview. Commun Assoc Inf Syst 19, 762–780.

Parker, C.J., 2012. A Human Factors Perspective On Volunteered Geographic Information. 2012. Loughborough University.

Peuquet, D.J., 1994. It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. Ann Assoc Am Geogr 84, 441–461.

Peuquet, D.J., 2005. Theme section on advances in spatio-temporal analysis and representation. ISPRS J Photogramm Remote Sens 60, 1–2.

Peuquet, D.J., Duan, N., 1995. An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. Int J Geogr Inf Syst 9, 7–24.

Pickles, J., 1995. Ground truth: the social implications of geographic information systems. Guilford Press.

Pickles, J., 2004. A History of Spaces: Cartographic reason, mapping, and the geo-coded wrold. Routledge, New York, NY, USA.

Pickles, J., 2005. Arguments, debates, and dialogues: the GIS-social theory debate and the concern for alternatives. In: Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W. (Eds.), Geographical Information Systems: Principles, Techniques, Management and Applications. Wiley, New York, pp. 49–60.

Poetz, M.K., Prügl, R., 2010. Crossing domain-specific boundaries in search of innovation exploring the potential of pyramiding. J Prod Innov Manag 27, 897–914.

Poetz, M.K., Schreier, M., 2012. The value of crowdsourcing: Can users really compete with professionals in generating new product ideas? J Prod Innov Manag 29, 245–256.

Polous, K., Mooney, P., Krisp, J.M., Meng, L., 2013. Mining Event-Related Knowledge from OpenStreetMap. In: Jukka M., K. (Ed.), Progress in Location-Based Services, Lecture Notes in Geoinformation and Cartography. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 109–124.

Pultar, E., Raubal, M., Cova, T.J., Goodchild, M.F., 2009. Dynamic GIS case studies: Wildfire evacuation and volunteered geographic information. Trans GIS 13, 85–104.

Pustejovsky, J., 1991. The syntax of event structure. Cognition 41, 47–81.

Raymond, E., 1999. The cathedral and the bazaar. Knowledge, Technol Policy 12, 23–49.

Rehrl, K., Gröechenig, S., Hochmair, H., Leitinger, S., Steinmann, R., Wagner, A., 2013. A Conceptual Model for Analyzing Contribution Patterns in the Context of VGI. In: Progress in Location-Based Services. pp. 373–388.

Reilly, T.O., 2005. What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software [WWW Document]. Commun Strateg. URL http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html?page=1 (accessed 5.1.15).

Rescher, N., 2008. Process Philosophy. In: Zalta, E.N. (Ed.), The Stanford Encyclopedia of Philosophy. The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, Stanford, CA.

Roberts, J.C., 2008. Coordinated Multiple Views for Exploratory GeoVisualization. In: Dodge, M., McDerby, M., Turner, M. (Eds.), Geographic Visualization: Concepts, Tools and Applications. John Wiley & Sons, pp. 25–48.

Rowe, G., Frewer, L.J., 2005. A Typology of Public Engagement Mechanisms. Sci Technol Human Values.

Salmon, W.C., 1984. Scientific Explanation and the Causal Structure of the World. Princeton University Press, Princeton, New Jersey.

Scharl, A., 2007. Media Platforms for Managing Geotagged Knowledge Repositories. In: Scharl, A., Tochtermann, K. (Eds.), The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 Are Shaping the Network Society. Springer London, London, UK, pp. 3–14.

Schlossberg, M., Shuford, E., 2005. Delineating “Public” and “Participation” in PPGIS. URISA J 16, 15–26.

Schuurman, N., 2006. Formalization matters: Critical GIS and ontology research. Ann Assoc Am Geogr 96, 726–739.

SciPy Developers, 2014. Scipy [WWW Document]. URL www.scipy.org

Seeger, C.J., 2008. The role of facilitated volunteered geographic information in the landscape planning and site design process. GeoJournal 72, 199–213.

Sewell, W.H.J., 2005. Logics of history: Social theory and social transformation. The University of Chicago Press, Chicago and London.

Sieber, R., 2007. Geoweb for Social Change. In: Workshop on Volunteered Geographic Information. Santa Barbara, CA, USA.

Sieber, R.E., 2006. Public Participation Geographic Information Systems: a literature review and framework. Ann Assoc Am Geogr 96, 491–507.

Silvertown, J., 2009. A new dawn for citizen science. Trends Ecol Evol 24, 467–471.

Slocum, T.A., McMaster, R.B., Kessler, F.C., Howard, H.H., 2008. Thematic Cartography and Geovisualization: International Edition, 3rd ed. Prentice Hall.

Smith, G., 2008. Tagging: people-powered metadata for the social web, 1st ed, Choice Reviews Online. New Riders Publishing, Thousand Oaks, CA, USA.

Snodgrass, R.T., 1995. The Tsql2 Temporal Query Language.

Sowa, J.F., 1999. Knowledge Representation: Logical, Philosophical, and Computational Foundations, First edit. ed. Brooks/Cole Publishing Co., Pacific Grove, CA.

SQLite, 2014. SQLite [WWW Document].

Stefanidis, A., Crooks, A., Radzikowski, J., 2013. Harvesting ambient geospatial information from social media feeds. GeoJournal 78, 319–338.

Steward, H., 1997. The Ontology of Mind: Events, Processes, and States. Clarendon Press, Oxford.

Subramanyam, R., Xia, M., 2008. Free/Libre Open Source Software development in developing and developed countries: A conceptual framework with an exploratory study. Decis Support Syst 46, 173–186.

Sui, D., Elwood, S., Goodchild, M., 2013. Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice, International Journal of Geographical Information Science. Springer Netherlands, Dordrecht.

Sui, D.Z., 2008. The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. Comput Environ Urban Syst 32, 1–5.

Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D., Kelling, S., 2009. eBird: A citizen-based bird observation network in the biological sciences. Biol Conserv 142, 2282–2292.

Sylvan, R., 1992. Process and Action: Relevant Theory and Logics. Stud Log 51, 379–437.

Tapscott, D., Williams, A.D., 2008. Wikinomics: How Mass Collaboration Changes Everything, Expanded. ed. Portfolio, Penguin Group, New York, New York, USA.

Tapscott, D., Williams, A.D.D., 2006. Wikinomics: How Mass Collaboration Changes Everything, Expanded. ed, Journal of Information Technology & Politics. Portfolio, United States, New York, USA.

Taylor, P.J., 1990. Editorial comment GKS. Polit Geogr Q.

Terwiesch, C., Xu, Y., 2008. Innovation Contests, Open Innovation, and Multiagent Problem Solving. Manage Sci.

Trompette, P., Chanal, V., Pelissier, C., 2008. Crowdsourcing as a way to access external knowledge for innovation: control, incentive and coordination in hybrid forms of innovation. In: 24 Th EGOS Colloquium. Amsterdam, pp. 1–29.

Tsou, M.H., 2003. An Intelligent Software Agent Architecture for Distributed Cartographic Knowledge Bases and Internet Mapping Services. In: Peterson, M. (Ed.), Maps and the Internet. Elsevier Press, pp. 229–243.

Tulloch, D.L., 2008. Is VGI participation? From vernal pools to video games. GeoJournal 72, 161–171.

Turkucu, A., 2008. Development of a conceptual framework for the analysis and the classification of "Public Participation GIS."

Turkucu, A., Roche, S., 2008. Classification fonctionnelle des Public Participation GIS. Rev Int géomatique Volume 18, 429–442.

Turner, A.J., 2006. Introduction to Neogeography, OReilly Short Cuts. O'Reilly Media.

Tutty, J.I., Martin, F., 2009. User Generated Design - Teaching and Learning with Web 2.0. In: Kidd, T.T., Chen, I. (Eds.), Wired for Learning: An Educators Guide to Web 2.0. Information Age Publishing, Charlote, NC, USA, pp. 43–60.

Tweddle, J.C., Robinson, L.D., Pocock, M.J.O., Roy, H.., 2012. Guide to citizen science: developing, implementing and evaluating citizen science to study biodiversity and the environment in the UK, Natural History Museum and Centre for Ecology and Hydrology.

Van Der Maaten, L.J.P., Hinton, G.E., 2008. Visualizing high-dimensional data using t-sne. J Mach Learn Res 9, 2579–2605.

Vendler, Z., 1967. Linguistics in Philosophy, ITHACA NY: CORNELL UNIV PR. Cornell University Press, Ithaca, NY.

Verona, G., Prandelli, E., Sawhney, M., 2006. Innovation and Virtual Environments: Towards Virtual Knowledge Brokers. Organ Stud 27, 765–788.

Von Kutschera, F., 1993. Sebastian Strolls. Grazer Philos Stud 45, 75–88.

Wagner, H.L., 1999. The psychobiology of human motivation. Psychology Press, London: Routledge.

Wahlberg, T.H., 2009. Objects in Time Studies of Persistence in B-time. Media-Tryck, Lund.

Walsh, R., 1992. Sociocultural perspectives on substance abuse disorders. Am J Psychiatry 149, 1760–1761.

Warren, D.M., 1991. Using Indigenous Knowledge in Agricultural Development; World Bank Discussion Paper No. 127. WASHINGTON, D.C.

Wasko, M.M., Faraj, S., 2000. “It is what one does”: why people participate and help others in electronic communities of practice. J Strateg Inf Syst 9, 155–173.

Weiner, D., Harris, T.M., 2003. Community-integrated GIS for Land Reform in South Africa. URISA J 15, 61–73.

Weiner, D., Harris, T.M., 2008. Participatory Geographic Information Systems. In: The Handbook of Geographic Information Science. Wiley-Blackwell, pp. 466–480.

Weiser, P., Frank, A.U., Abdalla, A., 2012. Process Composition and Process Reasoning over Multiple Levels of Detail. In: Geographic Information Science, 7th International Conference, GIScience 2012. Columbus, Ohio.

Whitla, P., 2009. Crowdsourcing and Its Application in Marketing Activities. Contemp Manag Res.

Wilson, J.P., Fotheringham, A.S., 2007. The Handbook of Geographic Information Science. Wiley-Blackwell, Malden, MA, USA.

Wilson, M.W., 2009. Towards a genealogy of qualitative GIS. In: Cope, M., Elwood, S. (Eds.), Qualitative GIS: A Mixed Methods Approach. SAGE, London. UK, pp. 156–170.

Worboys, M., 2005. Event-oriented approaches to geographic phenomena. Int J Geogr Inf Sci.

Worboys, M., Duckham, M., 2006. Monitoring qualitative spatiotemporal change for geosensor networks. Int J Geogr Inf Sci 20, 1087–1108.

Worboys, M., Hornsby, K., 2004. From objects to events: GEM, the geospatial event model. In: Egenhofer, M.J., Freksa, C., Miller, H.J. (Eds.), Geographic Information Science Lecture Notes in Computer Science. Springer, Berlin Heidelberg, pp. 327–343.

Worboys, M.F., 1994. Object-oriented approaches to geo-referenced information. Int J Geogr Inf Syst 8, 385–399.

Worboys, M.F., Hearnshaw, H.M., Maguire, D.J., 1990. Object-oriented data modelling for spatial databases. Int J Geogr Inf Syst 4, 369–383.

Wu, S.C., Fang, W., 2010. The effect of consumer-to-consumer interactions on idea generation in virtual brand community relationships. Technovation 30, 570–581.

Yuan, M., 2001. Representing Complex Geographic Phenomena in GIS. Cartogr Geogr Inf Sci 28, 83–96.

Zielstra, D., Zipf, A., 2010. A comparative study of proprietary geodata and volunteered geographic information for Germany. In: 13th AGILE International Conference on Geographic Information Science 2010. Guimarães, Portugal.

BIBLIOGRAPHY

Polous, K., Krisp, J.M., Meng, L., Shrestha, B., Xiao, J., 2015. OpenEventMap: A Volunteered Location-Based Service. Cartogr Int J Geogr Inf Geovisualization 50, 248–258.

Polous K., Freitag A., Krisp J., Meng L., 2015. A General Framework for Event Detection from Social Media. In Francis Harvey and Yee Leung (Eds.), Advances in Spatial Data Handling and Analysis (pp 85-105). Springer International Publishing.

Polous K., Mooney P., Krisp J., Meng L., 2013. Mining Event-related Knowledge from OpenStreetMap. In Jukka Krisp (Ed.), Progress in Location-Based Services (pp. 109-124). Springer Berlin Heidelberg.

Polous K., Krisp J., Meng L., 2012. The Note on Event for Event Detection from Volunteered Geographic Information (VGI), GIScience2012, Columbus, Ohio, USA. September 19-21, 2012.

Polous K., Krisp J., Fan H., Meng L., 2012. Location-based Event Detection from Open Street Map Data - A Case Study in Germany and China, The 20th International Conference on Geoinformatics 2012, Hong Kong, China. June 15-17, 2012.

Khatereh Polous

In this book, the concept of mapping goes beyond the principle of mapping an object as a conceptual geographic entity with a distinct spatial, temporal and attributive identity. The main goal is to present a conceptual model for managing geo-knowledge which handles real world dynamisms. It uses a generic event-oriented perspective to implicitly represent causal relationships among different components of a Spatio-Temporal Information System. From this new perspective, objects in space and time are considered merely as information elements of events, which are connected to other event elements.

ISBN 978-3-8439-2614-0

9 783843 926140