LEGAL ISSUES AND VALIDATION OF VOLUNTEERED GEOGRAPHIC INFORMATION

ANDRIY RAK

April 2013

TECHNICAL REPORT
NO. 283

LEGAL ISSUES AND VALIDATION OF VOLUNTEERED GEOGRAPHIC INFORMATION

Andriy Rak

Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

April 2013

PREFACE

This technical report is a reproduction of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering in the Department of Geodesy and Geomatics Engineering, April 2013. The research was co-supervised by Dr. David Coleman and Dr. Sue Nichols, and support was provided by the GEOIDE Network of Centres of Excellence, the Natural Sciences and Engineering Research Council of Canada, and the Province of Nova Scotia.

Rak, Andriy (2013). *Legal Issues and Validation of Volunteered Geographic Information.* M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical Report No. 283, University of New Brunswick, Fredericton, New Brunswick, Canada, 128 pp.

ABSTRACT

The Canadian Geospatial Data Infrastructure (CGDI) provides access to authoritative geographic datasets of Canada, which are the source of accurate and reliable data. The process of acquiring, updating and maintaining such datasets using traditional approaches, requires both time and costly resources. As a result, in many cases the datasets are out of date because of the high cost of maintenance. An alternative approach to reliably create and update authoritative datasets is linked to its integration with Volunteered Geographic Information (VGI). VGI provides a vast source of spatial information to government, industry and citizens. However, the integration of VGI with CGDI generates several questions, with VGI quality and legal issues at the forefront.

This research has investigated methods for assessing the quality of VGI, and describes the importance of a link between VGI and legal liability in the need for integration of VGI with CGDI. This research developed a prototype to validate data quality and examined legal liability issues around VGI to discover a strategy for possible integration of VGI with CGDI datasets. The research also provides four primary risk management techniques for CGDI to manage risks resulted from incorporating VGI into their datasets.

DEDICATION

I dedicate this research to my parents Mr. and Mrs. Rak and my sister Iryna Rak, for their interest in my education.

ACKNOWLEDGEMENTS

I gratefully acknowledge the following people without whom this research would not be possible:

- My special thanks to Dr. David Coleman and Dr. Sue Nichols my thesis supervisors for their feedback, research write-up, guidance and support throughout the period of study at the University of New Brunswick.
- Dr. Teresa Scassa and Dr. Hillary Young for their guidance on the legal part of this research.
- Mr. Titus Tienaah for his assistance during prototype development, thesis write-up, as well as providing advice on the research.
- Ms. Wendy Wells for her constructive feedback on thesis write-up.
- Sylvia Whitaker and Lorry Hunt who have always kindly assisted with all administrative issues.
- I am highly indebted to all the open source tools and projects used in this research.
- I am grateful for funding support for my research provided by GEOIDE Network of Centres of Excellence, the Natural Sciences and Engineering Research Council, and the Province of Nova Scotia.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AND – Automotive Navigation Data
API – Application Programming Interface
BOM – Buffer Overlay Method
CanVec – Canada Vector Database
CGDI – Canadian Geospatial Data Infrastructure
CMAS – Circular Map Accuracy Standards
EULA – End User License Agreement
GI – Geographic Information
GIS – Geographic Information System
GPS – Global Positioning System
HTTP – Hypertext Transfer Protocol
IP – Intellectual Property
LBS – Location Based Services
MVC – Model View Controller
NES – Notification and Edit Service
NMAS – National Map Accuracy Standards
NRN – National Road Network
NTDB – National Topographic Data Base
OGS – Open Geospatial Consortium
OSM – OpenStreetMap
RMT – Risk Management Technique
RMP – Risk Management Plan
SDI – Spatial Data Infrastructure
SDTS – Spatial Data Transfer Standard
TMS – Tile Map Service
UML – Unified Model Language
UNB – University of New Brunswick
URL – Uniform Resource Locator
USGS – US Geological Survey
VGI – Volunteered Geographic Information
WAF – Web Application Framework
WGS – World Geodetic System
WMS – Web Map Service

1. Introduction

Eventually, today's huge power stations and national
transmission grids might be superseded by a system that relies
on efficient local "micropower generators". The Economist, [1999]

As an analogy to Kuhn's [2007] comments in his article in *The Economist* [1999], "national transmission grids" of mapping agencies and other oligopolies of Geographic Information (GI) might one day be superseded by efficient local "microGI" generators. Now, a decade after the publication of this article, there is increasing research into what "microGI" generators using volunteered data mean for GIS science and GI markets.

In the literature, there are at least five terms defining the phenomena of geographic data collection and distribution by volunteers using the Internet. Terms such as *neogeography* [Turner, 2007; Sui, 2008], *ubiquitous cartography* [Gartner et al., 2007], *user-generated content* [Sieber, 2007], and *collaboratively contributed geographic information* [Bishr and Mantelas, 2008] have been proposed, but the most commonly used term in the literature is *volunteered GI* or *VGI* [Goodchild, 2007].

Although there is some disagreement on the definition of VGI, the "proliferation of user-generated content is viewed as a harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals" [Goodchild, 2007]. VGI is also seen as "an extension of critical and participatory approaches to geographic information systems" [Elwood, 2008] and a "by-product of the so-called 'Web 2.0' environment" [Flanagin and Metzger, 2008]. In this thesis, the term VGI will be used to describe GI which has been collected for the most part by non-professionals and submitted to a public or private GI database.

The process of mapping the Earth accurately was, until recently, the preserve of highly skilled, well-equipped, and organized individuals and groups. For many years, it was usually the role of surveyors, cartographers, and geographers to map the world and transcribe it on paper or, since the 1960s, into the computer. Less than a decade ago, it was common to assume that people needed a university-level degree to be able to measure the Earth and transcribe the information on paper or into the computer as well as expensive equipment and infrastructure to support their work.

[Haklay and Weber, 2008]

This, however, is starting to change through the introduction of VGI.

Two current successful examples of VGI are OpenStreetMap [OpenStreetMap, 2011a] and Wikimapia [Wikimapia, 2011a]. OpenStreetMap in March 2009 had only 79,325 users, buy by June, 2011 it already had 411,816 registered users [OpenStreetMap, 2011b]. Wikimapia in March 2009 had only 8,991,616 places described by users, but in 2011 it already had one million registered users and more than 14 million places added [Wikimapia, 2011b]. In both cases, the respective numbers of registered users and their contributions have grown rapidly over a relatively short period of time.

VGI began to develop rapidly after the emergence of Web 2.0 where ordinary citizens began to produce and share GI on the Internet. The trend increased after Google, Microsoft and Yahoo! made their web mapping Application Programming Interfaces (APIs) public [Goodchild, 2007; Rouse et al., 2007; Gould Carlson et al., 2008]. U.S. President Bill Clinton's announcement in May 2000 to remove the constraints of selective availability of the GPS signal [Clinton, 2000] was just as important for the development of VGI. This provided much improved accuracy for simple, low-cost GPS receivers. In practical terms, this made it possible to acquire the receiver's position with an accuracy of ± 6m to 10m in normal conditions, in contrast to roughly ± 100 metres before the "switch off" of selective availability [Haklay and Weber, 2008]. Such a decision plays an important role for VGI since, among technologies used to acquire GI, Location Based Services (LBS) are the ones most commonly used.

The interest in VGI has grown considerably over the past decade [Goodchild, 2008; Flanagin and Metzger, 2008; Bishr and Mantelas, 2008; Elwood, 2008]. There are predictions that data collected by ordinary individuals (volunteers) could be used by researchers, policy makers, citizen groups and private institutions for purposes such as:

- Monitoring environmental change [Elwood, 2008];
- Filling gaps in existing spatial databases [Elwood, 2008];
- Enhancing the information used by professionals and decision makers [Gouveia et al., 2004; Seeger, 2008]; and
- Possible significant source for geographers to understand surface of the Earth [Goodchild, 2007].

As noted by Budhatoki et al. [2008], the potential of users to supply GI is promising enough that researchers are now exploring the role of citizens in augmenting the means of GI collection: “the six billion humans constantly moving about the planet collectively possess an incredibly rich store of knowledge about the surface of the Earth and its properties” [Goodchild, 2007].

VGI importance is likely to grow in the future, taking into account its positive aspects of providing low cost, up-to-date geographic information. Coleman et al. [2009] noted that if: (1) volunteers are encouraged appropriately; (2) the processes are managed wisely; and (3) the potential and the limitations of their contributions are understood and used in the proper context, there is an opportunity “to produce and enjoy richer and more up-to-date databases than we have ever seen in the past”.

Due to the increasing importance and value that VGI is bringing into GIScience now, there is more research about issues with VGI. The main limitations of VGI as noted by Poser and Dransch [2010] are:

- Availability: "It is unknown *how much*, *which* and from *where* information will be supplied. Unlike a sensor network, information collection from the public cannot be planned in advance so as to yield an optimal configuration of observations for the phenomenon of interest".
- Data quality: "Unlike physical sensors, humans cannot be calibrated and do not comply with standards. Mostly, they are not trained for specific observations, and thus can intendedly or unintendedly provide false information".

This research investigates data quality aspects of VGI and designs a prototype for VGI validation to comply with authoritative datasets accuracy standards for their potential integration. Such integration of VGI with authoritative datasets also raises important legal considerations, such as: intellectual property issues, privacy, defamation, liability of users/producers [Janssen, 2010], and nature and motivations of volunteer contributors [Budhathoki et al., 2008; Coleman et al., 2009]. Liability is a primary issue that can deter organizations from incorporating VGI into their datasets. Due to the lack of research on this topic, some organizations consider it to be a better practice to exclude VGI as a viable option. In view of the benefits that VGI can bring, it is important to continue and deepen the research on liability concerns surrounding VGI so that organizations will not fear facing the legal liability risks that can arise and will be equipped with appropriate techniques to manage such risks. A part of this research investigates the liability effects of using VGI for integration with authoritative datasets applicable to Canada. Liability issues of using VGI are discussed by examining the liability in contract, as well as in tort according to Canadian common law.

VGI could become an important source of geographic data for authoritative datasets. The research reported in this thesis builds on earlier research conducted by MScE students of University of New Brunswick (UNB) Botshelo Sabone [2009] and Nyaladzani Jairo Nkhwanana [2009]. Its main purpose is to investigate ways in which we can validate VGI so that it can correspond to or be integrated with authoritative datasets standards. This research also highlights the importance of legal liability concerns surrounding VGI and its integration with authoritative datasets.

1.1 Research Background

Considerable research has been conducted on concepts, motivations, data quality and trust approaches with respect to VGI [Bishr and Kuhn, 2007; Goodchild, 2007; Bishr and Mantelas, 2008; Elwood, 2008; Flanagin and Metzger, 2008; Goodchild, 2008; Coleman et al., 2009; Sabone, 2009; Goodchild, 2009a; Haklay, 2010]. Less research has been devoted to legal aspects of VGI [Janssen, 2010; Onsrud, 2010; Rak et al., 2012].

Goodchild [2007] introduces the concept of VGI as a means of "using the Web to create, assemble, and disseminate geographic information". He presents the evolution of VGI and technologies that made VGI possible, and later describes the benefits and values of VGI by presenting "humans as sensors" who can sense data that no other sensors can.

Goodchild [2008] also brings up concerns regarding the validity of VGI, which in most cases is gathered by non-experts (volunteers) who do not have the required skills and knowledge when dealing with GI. Goodchild was also one of the first researchers who addressed the issue of open access to information produced by volunteers, which can compromise not only specific citizens but national security as well [Goodchild, 2008].

Some of these issues have been further investigated by Coleman et al. [2009], where the nature and motivations of volunteer contributors were identified. Volunteer contributors were described and characterized into five overlapping categories:

- “Neophytes”,
- “Interested Amateurs”,
- “Expert Amateurs”,
- “Expert Professionals”, and
- “Expert Authoritatives”.

Volunteer contributors were also characterized by Grossner et al., [2008] and Jones [2011]. Coleman et al., [2009] provided lessons learned from a review of volunteer contributors to on-line communities.

The quality of VGI is one of the most important issues in the research being carried out. Based on quality, VGI may or may not be used in the process of updating authoritative datasets. The quality of GI is measured as the difference between the data and the world they represent and becomes poorer as the data and corresponding reality diverge [Devillers and Jeansoulin, 2006; Maué and Schade, 2008]. Data quality can be defined as “fitness for purpose” [Chrisman, 1991; Coote and Rackham, 2008], in other words, how well does the data suit users’ requirements? In the world of VGI, quality control is difficult to achieve due to the “heterogeneous and dynamic nature as well as the volume of the VGI” [Sabone, 2009]. Some mechanisms for quality control have been established, some of them rely on “automated methods, others thorough peer reviews by specialists” [Maué and Schade, 2008].

This research presented here hypothesized that the positional aspects of VGI should be capable of meeting NTDB mapping specifications which state that 90% or more of all well-defined features in the dataset should be within ten metres of their true position. Similar research has been carried out outside Canada by Haklay [2010] who conducted a comparative study of OpenStreetMap and Ordnance Survey datasets. Results of the research showed that the accuracy of OpenStreetMap information is in Haklays words “of a reasonable accuracy” (approximately ± 6 m positionally different from data provided by Ordnance Survey) but, at the same time, this analysis highlighted the inconsistency of VGI in terms of its quality. Recent tests in the United Kingdom [Zulfiqar, 2008; Basiouka, 2009; Ather, 2009; Haklay, 2010], France [Girres and Touya, 2010], Germany [Zielstra and Zipf, 2010; Ludwig et al., 2011; Mondzech and Sester, 2011] and Greece [Kounadi, 2009] all demonstrated that the positional accuracy of OpenStreetMap data is comparable to corresponding traditional geographic datasets that are maintained by national mapping agencies and commercial providers.

Different approaches have been suggested for assessing the quality of VGI. Bishr and Kuhn [2007] suggest that because the quality of all geographic information is subjective to some extent, trust should be used as a proxy for VGI quality, and demonstrated in the form of user ratings. In other words they suggested that the quality of the VGI can be evaluated by assessing the credibility of the producers of the data. A distinction between the ‘‘accuracy’’ of information and its ‘‘credibility’’ thus arises. While “accurate information in most cases is likely to be perceived as credible, technically inaccurate information can also be perceived as credible as long as the information consumer believes it” [Flanagin and Metzger, 2008].

Bishr and Kuhn [2007] and Coleman et al. [2009] stated that, due to some definite spatial and temporal considerations that make VGI contributions unique, there are ways to validate contributions of spatial information and their contributors. Several recent studies [Budhathoki et al., 2008; McDougall, 2009] have agreed that the integration of VGI with official information provided by experts will be an innovative and powerful source of fresh data [Genovese and Roche, 2010]. Goodchild [2009a] has also commented that hybrid solutions to the production of GI may well represent the best of both worlds. Some recent examples of possible integration of VGI with authoritative datasets are:

- *US Geological Survey (USGS) tests out the social networking (Twitter) to collect real time, earthquake-related messages and get early response information about earthquakes. The USGS hopes to use this information to accelerate the delivery of risk information and earthquake response;*
- *Geomatics Canada evaluates the potential of OpenStreetMap to be part of the official national topographic database, and many others.*

[Genovese and Roche, 2010]

The main focus of ongoing research on legal aspects of VGI examines the legal ability to use GI, privacy, and security of VGI. Onsrud [2010] argues that there is a lack of an operational web-wide capability allowing users to legally access and use GI of others without seeking permission on a case-by-case basis, which in many cases takes a long time, and thus prevents many users from using available data. Sieber [2007] and Obermeyer [2007] debate about security issues of implementing VGI into authoritative databases noted that while some data is intentionally created and shared, other data may be collected from its ‘producers’ without their awareness or permission. Further research on legal issues associated with VGI is required, especially with respect to the purposes for which VGI may be used, ownership rights of the data, liability of both volunteer contributors and organizations using such data, and privacy concerns.

This section shapes the research problem: how can we validate the quality and address legal issues around VGI for it to be used as a source for updates to authoritative datasets. Through the literature review it has been identified that both of these two parts are necessary for a successful integration of VGI with authoritative datasets. It would be useful to have a prototype that enables the validation of data quality and legal issues around the use of VGI to update authoritative datasets.

1.2 Research Problem

The Canadian Geospatial Data Infrastructure (CGDI) provides access to authoritative geographic datasets of Canada, which are the source of accurate and reliable data. The process of acquiring, updating and maintaining such datasets using traditional approaches, requires both time and costly resources. As a result, in many cases such datasets are out of date because of the high cost of maintenance. An alternative approach to reliably create and update authoritative datasets is linked to its integration with VGI. Such integration of VGI with CGDI generates several questions, with VGI quality and legal issues at the forefront. VGI data quality and legal liability are among the primary issues that can deter CGDI and other organizations from incorporating VGI into their datasets. Due to the lack of research on this topic, some organizations consider it to be a better practice to exclude VGI as a viable option. If not properly managed, concerns over the quality and legal liability of VGI may limit its consideration before it reaches its full potential.

There are gaps in the research on actual implementation and testing of prototype tools which will be able to impose certain levels of trust and identify reliable content from VGI content. Obtaining reliable VGI is very important, as user-generated content can be more valuable (or in some cases, is *only* useful) when its spatial and temporal properties can be verified [Lenders et al., 2008]. No current studies in Canada have examined the legal liability arising from integration VGI with authoritative datasets, thus "…demonstrating the gap in the literature that whilst these technologies are being used …, little attention has thus far been paid to the potential legal liabilities arising from such use" [Low et al., 2010].

1.3 Research Objectives

The main purpose of this research is to validate VGI in accordance with CGDI standards. To achieve this goal, the following objectives have been set:

1. To design and deploy a prototype that is capable of validating VGI to comply with CGDI *positional accuracy* standards in particular.
2. To discuss and incorporate, where possible, in the prototype some of the legal liability concerns surrounding integration of VGI with CGDI datasets.

1.4 Methodology

To achieve the objectives of the research, there were two main steps. The first step was to discuss potential legal liability concerns surrounding VGI applicable to Canada. Legal liability was chosen among all other legal issues, because this topic has been mentioned but not greatly examined in the literature. Also legal researchers at University of Ottawa emphasized the usefulness of this topic in discussions with the author. To achieve above stated objective the literature review and its analysis have been done, and are described in Chapter 2. As part of the literature review, Canadian legislation was examined to discuss legal liability in contract and tort applicable to Canada, as well as literature on risk management techniques to limit/reduce potential legal liability. The literature review included only the principles of the common law, and did not cover the principles used in the Province of Quebec, which are governed by civil law. Discussion on legal liability concerns surrounding integration of VGI and CGDI provided in this research should be used to inform and raise the awareness around the issue; it is not a comprehensive legal review. The literature review was performed using materials from the following database engines (published in English language):

- CanLII,
- Google Scholar,
- HeinOnline,
- IEEE Xplore Digital Library,
- LegalTrac,
- LexisNexis,
- Quicklaw Plus,
- Scopus.

The second step addressed the design and development of a prototype - a web application, which is capable of providing tools to validate VGI with data from authoritative datasets. The validation of VGI was based on the following measures, described in Subsection 2.3: positional accuracy, attribute accuracy, data structure, metadata and uncertainty. The assumption made for this research is that an authoritative dataset is of high accuracy, i.e., represents the true position of features, and its temporal resolution is “good”, e.g., not older than 5 years. The prototype for validating VGI with data from authoritative datasets includes:

- A wiki web application that allows users to add and edit VGI;
- Enabling other users to provide feedback on the contributor and contribution;
- Incorporating risk management techniques to address legal liability concerns surrounding integration of VGI and CGDI; and

- A validation engine which serves as the backend of the web application.

The prototype has three main components:

- Web client application,
- Web and spatial server, and
- Spatial database server, which will hold both authoritative dataset and VGI.

Python and JavaScript were used to script server and client side of the prototype. Python and JavaScript were chosen as they are among the most common programming languages used to develop web applications. PostgreSQL, an open source object-relational database system, was used as a spatial database server. PostgreSQL was chosen for this research, as it runs procedures from Python programming language, and it supports “PostGIS” extension, which adds support for geographic objects in PostgreSQL, allowing it to be used as a spatial database for Geographic Information System (GIS) [PostgreSQL, 2011]. PostgreSQL is compliant with the Open Geospatial Consortium (OGS) standards, thus providing a consistent programing interface between prototype components.

The prototype was used as a medium to validate VGI with data derived from the GeoBase portal (part of CGDI). VGI validated in this research came from field measurements collected by Sabone [2009], and represents residential streets in Skyline Acres, a subdivision within the city of Fredericton, New Brunswick, Canada, and is described in Subsection 4.2. To validate VGI, an authoritative dataset for the City of Fredericton, New Brunswick, Canada from GeoBase portal was used. The prototype itself was not tested, only VGI by comparing it with data from GeoBase portal.

The prototype developed in this research would be deemed successful if it satisfies the criteria below:

- Satisfaction of the CGDI accuracy standards (described in Subsection 2.3) by VGI after being validated through the prototype; and
- Incorporation of tools (described in Subsection 3.4.7) within the prototype to limit/reduce potential legal liability of GI providers from incorporating VGI into their datasets.

The prototype was used to validate VGI by both graduate and undergraduate students from the University of New Brunswick, Canada, as well as by other users interested in the research, and is discussed in Subsection 4.1.

The research methodology is diagrammed in Figure 1.1. The validation of VGI and evaluation of the prototype are discussed in Chapter 4.

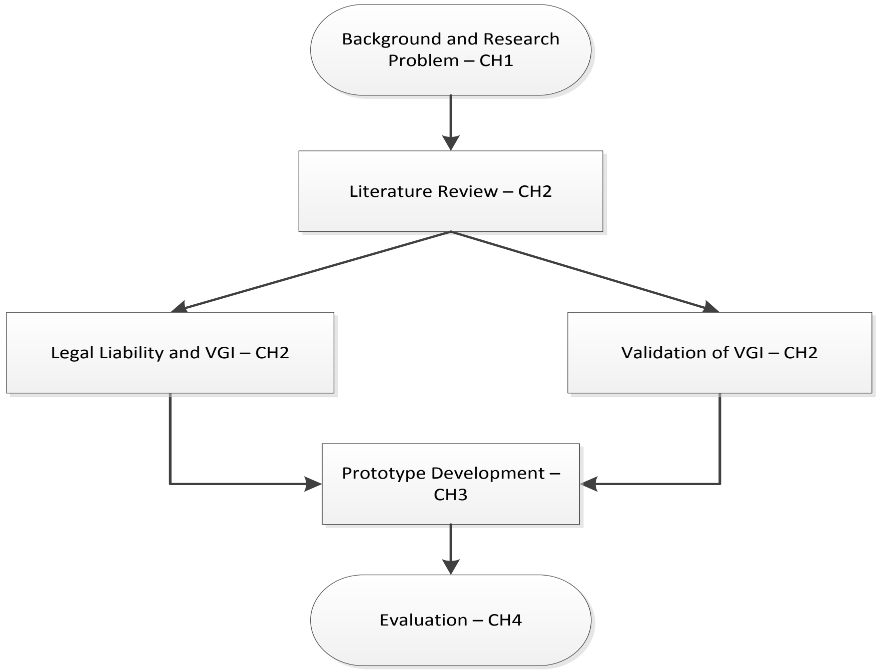


Figure 1.1 Research Method

1.5 Research Contribution

The importance of integrating authoritative datasets with VGI has been acknowledged by Budhathoki et al. [2008], Coleman et al. [2009], Goodchild [2009a], McDougall [2009] and Genovese and Roche [2010]. This research describes validation of VGI to comply with CGDI datasets accuracy standards, described in Subsection 2.3 that can lead to integration of VGI and CGDI datasets. In addition to the technological development of a prototype, this research also discussed legal liability aspects around VGI. This research helps to fill the knowledge gap that currently exists due to lack of:

- Study of legal issues arising from the implementation of VGI with authoritative datasets;
- Practical solutions to validation of VGI form technological perspective, by developing a prototype.

1.6 Organization of Thesis

This thesis is organized into five chapters. This chapter introduced the research problem, the key objectives, and discusses the significance of this research. Chapter 2 is a literature review introducing VGI, different ways of VGI validation to comply with authoritative datasets standards, and the quality of VGI contributions. Chapter 2 also includes a literature review and discussion of some of the legal liability issues arising from integrating VGI with data from authoritative datasets.

Chapter 3 discusses the design and prototype implementation for validating VGI with CGDI datasets. This chapter also discusses system architecture and processes of interaction among client, server, and database.

Chapter 4 presents an evaluation of the prototype in comparison to criteria set in the thesis objectives. This chapter discusses the role of the prototype in validation of VGI, and concludes with the research outcome.

Chapter 5 summarizes this research with recommendations for future work.

2. Data Quality of VGI and Legal Liability Issues

In this Chapter, the concepts of VGI and GI quality are discussed. This chapter outlines and defines validation techniques used by both authoritative organizations and open source projects in assessing the quality of their datasets. A review of CGDI accuracy standards, which have to be met for a VGI to be used as a source of updates for CGDI datasets, is also included.

The chapter also describes the importance of the link between VGI and legal liability in the need of integration of VGI with CGDI. The questions of who is liable and under what circumstances if VGI is provided to authoritative public and private geographic datasets are among the most important questions that impact VGI. Liability issues of using VGI are studied by examining the liability in contract, as well as tort according to Canadian common law system. Discussion on legal liability concerns provided in this Chapter should be used to raise the awareness around the issue, and not to provide a comprehensive legal review. The chapter concludes with a summary of the main risk management techniques that are required to be developed by an organization to minimize and/or eliminate legal liability when using VGI.

2.1 VGI and Geographic Information Quality

The quality of GI is a developing concept. The quality of the data defines the value of the data and, therefore plays an important role in determining whether such data can be used for a specific purpose. There are two primary concepts of GI quality. The first concept defines the quality of GI as "a measure of the difference between the data and reality that they represent, and become poorer as the data and the corresponding reality diverge" [Devillers and Jeansoulin, 2006], and is referred to as an internal data quality in literature.

The second concept describing data quality is based on a "fitness for use" concept, also referenced to as an external quality. This concept is becoming widely used among organizations exploiting GI, and the literature states that "… ultimately it is the consumer who will judge whether or not a product is fit for it to use" [Chrisman, 1983; Wang, 1996; Devillers and Jeansoulin, 2006; Coote and Rackham, 2008]. One of the important characteristics of this concept is that "…the producer responsible for the test, the evaluation [on data], and the assessment, … do not all necessarily belong to the same group of people" [Devillers and Jeansoulin, 2006]. As an example how to assess quality of VGI, the following steps can be taken:

1. GI tested for a "quality check" goes through all of procedures required by the specific "quality check" requirements; a report on the satisfaction of the data quality check requirements is then generated, which is then published/distributed together with the data.

2. The users will later decide, based upon the report, whether such data satisfies their needs.

External data quality is a relative measure: there are no absolutes. In assessing any product or service, whether it is a piece of furniture or a compiled database of GI, quality "… is meaningless unless it is expressed as a measure against a production specification or user requirement" [Coote and Rackham, 2008]. Data quality and quality checks with respect to VGI differ slightly from authoritative datasets created/updated using traditional methods. Although there is no official definition of the term "authoritative dataset", the term "Authoritative Data" has been used to describe products produced by professional mapping organizations [Goodchild, 2009b; Ball, 2010; Coleman et al., 2010], and are typically "… considered authoritative because they originate with a reputable source (e.g., a government or established industry actor) or because they meet measureable standards for quality and credibility" [Scassa, 2012]. Among the main perceived differences between VGI and authoritative datasets is the expectation of internal data quality. For authoritative datasets, there are reasonable expectations that a highly trained crew, using precise measurement equipment, trained in the field of data capturing, would provide "highly" accurate data within their field of work. With regards to VGI, such expectations are not always so convincing.

Another difference is the absence, in most cases, of metadata for VGI, and its presence for authoritative datasets. Metadata is a file with information "… which captures the basic characteristics of a data or information resource. It represents the who, what, when, where, why and how of the resource" (Federal Geographic Data Committee [FGDC, 2012]). Due to these factors and depending on techniques used to validate VGI quality, it is difficult for organizations that create and/or employ GI in their business, to trust and as a result utilize VGI. Although it should be mentioned that such resistance by organizations to employ VGI, lastly is beginning to diminish, as new studies [Zulfiqar, 2008; Ather, 2009; Basiouka, 2009; Sabone, 2009; Haklay, 2010; Zielstra and Zipf, 2010; Mondzech and Sester, 2011] have proven that VGI has the potential to be one of the sources of GI for authoritative datasets.

2.2 VGI Validation

Quality control measures are performed on VGI, to ensure that "non-conformance with data specifications is either prevented from occurring or identified and corrected as close to the source as possible" [Devillers and Jeansoulin, 2006]. Quality of VGI, as described in section 2.1, is very important, as "user generated content can be more valuable (or in some cases, is *only* useful) when its spatial and temporal properties can be verified" [Lenders et al., 2008].

Approaches towards validation of VGI can be grouped into two main groups:

- Traditional/authoritative, and
- Crowdsourcing.

2.2.1 Authoritative Methods of VGI Validation

Although there is no official definition of the term "authoritative" methods of assessing the quality of GI, in this research the term has been used to describe methods for assessing quality of GI that are based on a well-established, proved techniques that provide reliable and confident of quality GI.

Different approaches have been used by authoritative organizations to assess the quality of GI, which also can be used for assessing the quality of VGI. Metadata, as one of the approaches, has been used to "facilitate accessing up-to-date and high quality spatial data and services" [Williamson et al., 2003; Rajabifard et al., 2009], and has been an important part of SDI efforts for over 20 years now. Likewise such approaches as the US National Map Accuracy Standards (NMAS) and Circular Map Accuracy Standards (CMAS) can be used to assess the positional accuracy of GI. However, with the development of data quality concepts, such as "fitness for use" and "accuracy beyond position", it became clear that the original producer of the data "… cannot foresee all potential users. It became crucial for a data producer to record important aspects about the data so that users may make informed judgments regarding fitness for purpose" [Devillers and Jeansoulin, 2006]. Thus the Spatial Data Transfer Standard (SDTS) was developed, which provides "a data-quality report containing five parts … lineage, positional accuracy, attribute accuracy, logical consistency, and completeness" [Rossmeissl, 1989; Devillers and Jeansoulin, 2006] and identifies four methods for assessing the positional accuracy of a digital dataset:

- Deductive estimate (practical estimate of errors in the source of spatial data including assumptions made about error propagations);
- Internal evidence (all possible statistics or adjustments that may be used on GI);
- Comparison to source (comparing the derived spatial data with the original source);
- Comparison to source of higher accuracy, which had been suggested as a preferred method for assessing positional accuracy of a digital dataset [USGS, 1999; Goeman et al., 2005].

While standard practice for checking the positional accuracy of GI still relies on comparison of the points, current “applications put stricter demands on shape fidelity and relative accuracy” [Goeman et al., 2005]. This is particularly important for VGI, as a vast source of such information is coming in the form of linear features recorded using handheld GPSs and GPS enabled phones. The most commonly used method for assessing positional accuracy of VGI is the Buffer Overlay Method (BOM). The BOM introduced by Goodchild and Hunter [Goodchild and Hunter, 1997], indicates in percentage terms, the data set length included within a specific buffer distance from the reference data set. The BOM principle is shown in Figure 2.1.

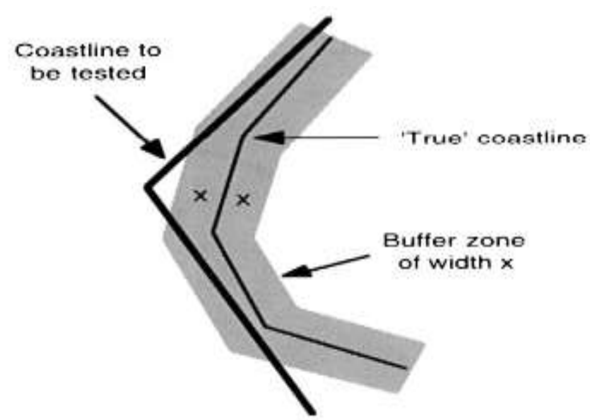


Figure 2.1 Buffer Overlay Method (from Goodchild [1997])

BOM has been used in the most recent studies on VGI quality carried out by:

- Sabone [2009] to test quality of VGI against CGDI accuracy standards;
- Haklay [2010] to test OpenStreetMap and Ordnance Survey datasets for Great Britain;
- Girres and Touya [2010] to test OpenStreetMap datasets for France;
- Kounadi [2009] to test OpenStreetMap datasets for Greece.

As such BOM is seen as the most popular method for validating internal quality (positional accuracy) of VGI, and was used in this research. Other methods for checking VGIs both internal and external quality may be adopted later within the prototype.

2.2.2 Crowdsourcing Methods of VGI Validation

Crowdsourcing is a relatively recent concept and covers a number of activities.

> *Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage that what the user has brought to the venture, whose form will depend on the type of activity undertaken.*
>
> [Estellés-Arolas and González-Ladrón-De-Guevara, 2012]

The successes of leading e-Commerce projects such as *eBay.com* and others could hold a key to building crowdsourcing approaches for evaluating VGI quality, for example by using a 5-star rating system for 3D VGI [Jones, 2011]. Different lessons can be learned from the existing handling of e-Commerce initiatives and also from leading wikis such as *Wikipedia.org* [Nkhwanana, 2009]. Wikipedia, one of the first wiki projects, validated the information in their database by relying on the wisdom of crowds, a type of crowdsourcing activity. The main argument was that content created by many is more likely to be accurate than content created by a few [Surowiecki, 2005]. However, the recent and growing problem of "vandalism" on the site made Wikipedia change its trust system approach. "We are no longer at the point that it is acceptable to throw things at the wall and see what sticks," mentioned Michael Snow, Chairman of the Wikipedia Board [Beaumont, 2009]. The new approach employs "flagged revisions", which enables editors to adjudicate on any changes made to the Wikipedia pages of a living person before those changes go live.

It can be seen from lessons learned from Wikipedia that, at the stage when databases do not have large numbers of users and contributions, the issues of data quality did not have priority. However, as soon as there became considerable amount of data within the Wikipedia or OpenStreetMap projects, stricter measurements were implemented to check and maintain data quality. The idea for stricter trust approaches could be considered when designing a prototype for implementing VGI with authoritative databases. Most authoritative databases have quality assurance rules and VGI contributions should not degrade the quality of authoritative databases, but rather enrich them.

Bishr and Kuhn [2007] suggest that "because the quality of all geographic information is subjective to some extent, trust should be used as a proxy for VGI quality, and demonstrated in the form of user ratings", which are very subjective but could very well be an indicator of data quality. In other words, it was suggested that the quality of

VGI can be evaluated by assessing the credibility of data producers. A distinction between the ‘‘accuracy’’ of information and its ‘‘credibility’’ thus arises. “While accurate information in most cases is likely to be perceived as credible, technically inaccurate information can also be perceived as credible as long as the consumer believes in the information” [Flanagin and Metzger, 2008].

As Goodchild [2009a] noted, “… both traditional/authoritative and crowdsourcing approaches provide inaccurate data, but the inaccuracies of data traditionally produced are tested and well documented” [Delavar and Devillers, 2010]. VGI, for the most part, is recorded in a digital vector format, where the real world is represented using points, lines and polygons (OpenStreetMap, Wikimapia, Tagzania, The People’s), with some information being annotated using pictures and text (The People’s Atlas, Wayfaring). Consequently, most quality controls performed on VGI are directed towards quality controls on vector data.

2.3 VGI and Canadian Geospatial Data Infrastructure Accuracy Standards

The CGDI initiative is an inter-governmental effort, which “encompasses the standards, technology, access systems and institutional arrangements” [GeoConnections, 2012b] to “improve the sharing, access and use of Canadian geospatial information” [GeoConnections, 2012a]. CGDI would benefit from the integration of VGI into their databases. “VGI could enhance the spatial data used in SDI’s, especially when it comes to data that is lacking or in need of updating" [Sabone, 2009]. One of the projects integrating VGI into CGDI was carried out by Geomatics Canada, which evaluates the potential of OpenStreetMap to be part of the official national topographic database [Genovese and Roche, 2010].

Such potential integration of VGI and CGDI raises several questions on the relationships between VGI and Spatial Data Infrastructures (or "SDIs"). Castelein et al. [2010] suggests that there has been little empirical investigation analyzing VGI phenomena and the convergence of VGI and SDI. Bishr and Mantelas [2008] argue that VGI is unique in comparison to SDI as it is based on a bottom-up approach in comparison to top-down approach in SDI. Goodchild [2008], on the other hand, argues that VGI fits in the model of SDI because it facilitates exchange of geographic information between individuals in a community. However, as McDougall [2009] notes, a proper validation framework, as well as the means for verifying contributors' expertise and credibility, must first be established before integrating VGI into SDI. In his opinion, if this validation framework is not in place, VGI cannot be part of SDI.

In earlier research by Sabone [2009], criteria and specifications for VGI to be integrated with GSDI were derived from the National Topographic Data Base (NTDB), National Road Network (NRN), and Canada Vector Database (CanVec). They were analyzed and grouped together into five requirements that employed the specifications mentioned above to assess VGI and compare with the CGDI datasets. Refer below for the list of criteria:

1. ***Data Structure**: The data should have entities that have a name to distinguish it from other entities or features, such as a road or bus stop,*

and the entities should comprise geometric representation, descriptive representation, an identifier and metadata (NRN and NTDB specifications).

2. ***Metadata**: The metadata for a dataset or feature at the minimum should include: source of data (e.g. GPS) and accuracy of data during acquisition.*
3. ***Positional Accuracy:** According to the respective NTDB, CanVec and NRN data specification documents, VGI must have a planimetric accuracy of no less than ±10 m for roads (in clear unobscured areas), and ± 30 m or less for other features (in clear unobscured areas) in order to be integrated and used with these datasets.*
4. ***Attribute Accuracy:** Attributes should include a name or identifier, e.g., road name for a road; and descriptive information, such as type or function.*
5. ***Uncertainty:** The maximum acceptable proportion of errors of classification for attributes should be no more than 5% and the maximum acceptable proportion of positional errors should be no more than 10%. The criterion examines the errors that exist within the VGI datasets as a whole. The clause for attribute uncertainty is based on NRN and CanVec specifications; while the clause for positional uncertainty is based on traditional mapping specifications which state that 90% or more of all well-defined features in a dataset should be within a specified distance (in this case ± 10 m) of their true position*

[Sabone, 2009].

Sabone [2009] discussed requirements for integrating VGI with CGDI based on the above mentioned five criteria for VGI; however, she did not include legal considerations arising from the use of VGI in the process of updating authoritative datasets. Due to the importance of legal liability in the potential use or integration of VGI and CGDI, the need to consider legal considerations is required. Requirements proposed for integrating VGI and CGDI need to be updated and should include the legal considerations. This research proposed that an additional criterion should be established:

6. ***Legal Considerations***: For VGI to be integrated with other datasets of CGDI, it must correspond not only to data quality standards, but also to legal rules and regulations including, for example, intellectual property, privacy, and legal liability.

2.4 Liability Concepts and VGI Quality

Both VGI and CGDI datasets are subject to errors. Utilization of VGI to maintain CGDI datasets brings positive aspects into the updating process, as mentioned in Chapter 1. However, VGI can bring some additional uncertainties into CGDI datasets. As a result, several potential legal liability issues might arise in the case when an inaccurate GI has been disseminated by CGDI. This section investigates how VGI may cause physical injury and result in liability. In Sections 2.5 and 2.6 this potential liability for physical injury will be explained under liability in contract and tort.

The risks for organizations employing Web 2.0 technologies and VGI has been acknowledged by Low et al. [2010]. They mention not only the opportunities that Web 2.0 and crowdsourcing can bring to the process "of collecting, obtaining and disseminating real-time [crisis] information to the public", but also stress the importance and limited amount of research on the potential legal liability from using VGI and Web 2.0 technologies.

A key concern originating from the integration of VGI and CGDI relates to the dissemination and use of inaccurate information introduced by VGI. The potential integration of VGI with CGDI datasets presents difficulties because maintaining accuracy of CGDI datasets would depend on the accuracy of VGI collected by volunteers. "Accordingly, the incorporation of user-generated information within official government information collection structures could mean that inadvertent, but nonetheless inaccurate, information is published thus opening the government agency to potential legal actions" [Low et al., 2010].

A key question in liability with respect to VGI concerns whether or not data is suitable for the specific use [Chandler and Levitt, 2011, p.85] - the so-called principle of "fitness for use" [Devillers and Jeansoulin, 2006]. If the users/consumers of VGI employ the data, and the data satisfies their needs, then it is likely that liability issues will not emerge. That outcome might be different if the data employed did not satisfy user needs and/or caused harm. An important question with respect to fitness for use is "… who bears the responsibility for assessing it?" [Chandler and Levitt, 2011, p.85-86]. Organizations that employ VGI take it for granted that such evaluation is within the scope of user responsibility [Gervais, 2006; Chandler and Levitt, 2011]. However, this is not always the case. For example, *Ontario's Consumer Protection Act* [2002] states that "any attempt to contract out of statutory implied conditions and warranties applicable to goods and services provided to consumers is void" [Chandler and Levitt, 2011, p.85-86]. The intention is that if there were or would be "statutory implied conditions" stating that an organization employing VGI is required to validate it before disseminating, such responsibilities could not be contracted out by that organization and transferred to a user.

The following section describes two cases where reliance on maps or other spatial data have caused physical injuries and/or loss of life. These cases show that usage of GI, including VGI as possible source of updates, can cause significant damages. In such circumstances legal cases are likely to arise, and with that the questions of who is liable for negative outcome.

2.4.1 Examples of Legal Liability Cases Based on Use of GI

Case 1: In Australia, four New South Wales National Parks and Wildlife Service (NPWS) officers died in the operation to put down a fire. During the operation officers used maps with two possible escape routes, neither of which in reality was suitable: one route had impenetrable bushes and the other a line of cliffs [Cho, 2005]. "A conclusion of the Deputy State Coroner was that the original botanical map had not been ground-truthed to include specific details and did not mark areas with safe refuges to retreat to as required in the fire management procedures guidelines" [Brown, 2001].

Case 2: Four men at Gretley Colliery, Newcastle Wallsend Coal Company Pty Ltd in New South Wales (N.S.W.), Australia, drowned due to unexpected water that came from a long abandoned old working of the Young Wallsend Colliery. A plan that miners used was approved by the N.S.W. Department of Mineral Resources (DMR). When the matter was investigated it was discovered that the department had made the original mistake of creating the incorrect maps — "the 'fatal error' that sparked the tragic chain of events" [Cho, 2005].

From the cases described above it can be noted that authoritative datasets, and VGI as one of the potential sources of updates to such datasets, could lead to unintentional circumstances such as physical injuries and can even cause loss of life. If the recipient of information makes a successful claim against CGDI for example on the basis that GI from their datasets contained errors, and caused damages, the CGDI could be potentially liable for physical injuries.

2.5 Liability in Contract

Currently CGDI is not utilizing VGI as a source of updates to their datasets, although this topic is being discussed. On November 18th, 2008 an agreement was made between GeoConnections (Natural Resources Canada) and OSM, to contribute geospatial datasets from the GeoBase initiative to the OSM foundation. It was proposed "that the GeoBase data be included in OSM based on the license compatibility and the aligned GeoBase and OpenStreetMap goals of providing free geospatial data to as wide audience as possible" [Wikipedia, 2008]. Such provision of authoritative datasets to one of the biggest VGI projects can also potentially lead to a reverse process, where VGI data would be used as one of the sources to update CGDI datasets, as Geomatics Canada evaluates the potential of OpenStreetMap to be part of the official national topographic database [Genovese and Roche, 2010]. But such an introduction of VGI into the CGDI database could potentially include some errors, and as a result a question of legal liability of CGDI could arise.

Legal liability for CGDI arising from the integration of VGI can be raised either in the Law of Contracts or in the Law of Torts. This section describes the fundamentals of legal liability in contract, and Section 2.6 will provide insights into legal liability in torts. The purpose in Sections 2.5 and 2.6 is to highlight main areas that will require future research.

Liabilities in contract are "claims based on breach of contract or warranty including implied warranties at common law, collateral warranties, and other statutory conditions" [Edgell, 2000]. The contracts regarding VGI relationships can cover many aspects, but this section will cover only the general liability principles concerning VGI and liability in contract, this includes:

- Liability under the contract;
- Parties to the contract.

Liability arises when the norms stated in the contract are broken by any of the parties to the contract. Contractual relations with respect to VGI are discussed between the following parties:

- The contributor of VGI and CGDI;
- CGDI and users/recipients of the information.

Current practice shows that contracts between the volunteer contributor and GI providers are not typically concluded. GI providers in their relationships with volunteers contributing VGI, have no practical need for a contract. Volunteers contribute VGI based on free will and have no obligation to contribute such data. They can continue contributing or stop when they wish. Liability in contract does not typically arise in this relationship, since there is no consideration from the volunteers to contribute or verify the quality of VGI.

In the case of the second type of relationships - i.e., between CGDI and users/recipients of GI - the extent of liability to which CGDI as provider is subject is unclear. Parties to the contract can agree on their own set of rules, including liability questions, unless those rules contradict the current law. Organizations that employ VGI in their business processes intend to limit their contractual and tort liability, by including waivers of liability in regards to data usage. For example, GI providers can include in a contract that data is used at the sole risk of the user, and the organization disclaims any liabilities for physical harm caused by data usage, or any products or services derived from such data. GeoBase [2012b] portal, for example, indicates the following concerning legal liability:

> *Canada makes no representation or warranty of any kind with respect to the accuracy, usefulness, novelty, validity, scope, completeness or currency of the Data and expressly disclaims any implied warranty of merchantability or fitness for a particular purpose of the Data. Canada does not ensure or warrant compatibility with past, current or future versions of your browser to access the site's Data.*
>
> *The Licensee shall have no recourse against Canada, whether by way of any suit or action, for any loss, liability, damage or cost that the Licensee may suffer or incur at any time, by reason of the Licensee's possession or use of the Data.*
>
> [GeoBase, 2012a]

While it is true that, in some cases, this could limit the potential legal liability, in other cases it will not be possible, as "consumer protection legislation may limit the enforceability of attempts to disclaim liability" [Chandler and Levitt, 2011, p.91].

2.6 Liability in Tort

"Tort Law consists of a body of rights, obligations, and remedies that is applied by courts in civil proceedings to provide relief for persons who have suffered harm from the wrongful acts of others" [Free Dictionary, 2011]. Tort law "determines when the person who causes the harm must pay compensation to the person who suffers it" [Osborne, 2007]. The principles discussed in this Section are those of the Canadian common law, and as such would not necessarily cover the civil law principles used in the Province of Quebec.

This thesis, follows the precedent established in *Sea Farm Canada Inc. v. Denton* [1991] and discusses the liability issues coming from inaccuracies in VGI under the rules of product liability, which is part of the law of negligence, and is also a sub-field of tort law [Chandler and Levitt, 2011], and is shown in Figure 2.2.

Figure 2.2 Tort Law in Canada

Black's Law Dictionary includes the following definition for negligence:

> *Negligence is the failure to use such care as a reasonably prudent and careful person would use under similar circumstances: it is the doing of some act which a person of ordinary prudence would not have done under similar*

> *circumstances or failure to do what a person of ordinary prudence would have done under similar circumstances.*
>
> [Black, 1979]

Tort law rests the burden of proof for negligence on the plaintiff's shoulders, and he/she must show "only a 51 percent likelihood that the facts of the case are as he/she alleges" [Osborne, 2007].

The person who has suffered physical harm from using VGI as a part of CGDI datasets must be able to prove the following elements to obtain a judgment of negligence [Osborne, 2007]:

- Duty of Care,
- Standard of Care,
- Causation,
- Remoteness of Damage, and
- Damage.

The first element of negligence that the plaintiff has to prove is the duty of care. The establishment of a duty of care "is a question of law which requires the judge to determine if the defendant is under a legal obligation to exercise reasonable care in favor of the plaintiff" [Osborne, 2007]. In the *Donoghue v. Stevenson* [1932] case a *neighbour* principle was established to "formulate a general conception of relationships which gives rise to a duty of care" [Osborne, 2007]. It was described as:

> *The rule that you are to love your neighbour becomes in law, you must not injure your neighbour; and the lawyer's question, Who is my neighbour? receives a restricted reply. You must take reasonable care to avoid acts or omissions which you can reasonably foresee would be likely to injure your neighbour. Who, then, in law is my neighbour? The answer seems to be – persons who are so closely and directly affected by my act that I ought*

> *reasonably to have them in contemplation as being so affected when I am directing my mind to the acts or omissions which are called in question.*
> [*Donoghue v. Stevenson*, 1932]

It is common ground that the test for determining whether the duty of care exists in a particular situation has been set in *Cooper v. Hobart* [2001]. The test outlined the following steps:

- "First it is necessary to decide if the case falls within any recognized category of relationships where a duty of care has previously been recognized or if it is so closely analogous to a recognized category. In those circumstances, a prima facie duty of care may be posited" [Osborne, 2007];
- If the case falls outside an established category a three-stage test for the 'novel' case should be used, where foreseeability, proximity, and residual policy factors should be established [Osborne, 2007].

To establish that a duty of care has been owed by CGDI to its users, it is unnecessary to foresee the identity of particular victims, but just to prove that a certain class of victims would be harmed by CGDI carelessness. In a case where CGDI provides a careless GI updated using VGI, at least certain kinds of harms will be foreseeable. As such, CGDI, being led by the Natural Resources Canada, a ministry of the government of Canada, could be recognized to owe a duty of care to the plaintiff for the harm caused by VGI. However, if the decision of CGDI to use VGI for budget reasons will be interpreted as a policy decision, there will be no duty of care, as government does not have a duty of care to the plaintiffs for its policy decisions [Osborne, 2007].

The primary element "of negligence liability is the negligent act, a failure to take care for the safety of the plaintiff. In determining the appropriate degree of care, it is useful to have some standard against which to measure the conduct of the defendant" [Osborne, 2007]. The common law uses the standard of a *reasonable person* when discussing the standard of care element in negligence, and focuses on whether the defendant fell below that standard [Osborne, 2007]. The standard of a reasonable person is an objective criterion and describes what a reasonably prudent person under the same or similar circumstances would do. The standard of conduct "is determined by taking into account both the practical realities of what ordinary people do and what judges believe they ought to do. It is not, however, the standard of perfection. The reasonable person may make mistakes and errors of judgment for which there is no liability" [Osborne, 2007].

> *When applying the standard of care to the facts of a case, courts have found it useful to review a number of factors that guide a reasonable person in the regulation of her conduct. The most significant of these is the concept of a reasonably foreseeable risk. Conduct is negligent only if it carries a risk of damage that a reasonable person would contemplate and guard against. In addition, courts routinely consider the likelihood of damage occurring, the seriousness of the threatened harm, the cost of preventive measures, the utility of the defendants conduct, any circumstances of emergency compliance by the defendant with approved practice or custom, and the defendant's post-accident precautions. Other more controversial and elusive factors such as judicial policy, economic factors, the equity of the individual case, and the psychological phenomenon of hindsight bias may also be influential.*
>
> [Osborne, 2007]

In case when a defendant is a professional or someone with special skills, the standard of a reasonable person also depends on the assumed knowledge and experience of the defendant [WikiBooks, 2013], and can be raised [Osborne, 2007]. For example, professionals would be judged according to the standard of care of other professionals in the same or similar field of work [Osborne, 2007]. The standard of care for CGDI, being a source of an authoritative data for entire Canada, would be that of a prudent and competent member in the field of geomatics in Canada.

The main issue with respect to establishing standard of care for the CGDI would be the issue of reasonableness, specifically whether use of VGI as a source of updates for CGDI datasets is reasonable (no liability) or unreasonable (liability). In differentiating between these two situations "courts must be sensitive to current attitudes to risk and the importance of facilitating the legitimate activities, needs, and aspirations of citizens… Many foreseeable risks must be tolerated because they are integral to the kind of society that we have chosen to create and live in" [Osborne, 2007]. To give an answer to this question a cost-benefit analysis accompanied by assessment of "all the relevant considerations" [Osborne, 2007] would take place to find out if the defendant acted reasonably, particularly would the benefits that VGI brings outweighs the possible harm and damages.

The question of whether there would have been a breach of standard of care by CGDI in a case of using VGI as a source of updates would be decided by a judge or by a jury. Some of the considerations that should be taken into account while performing a cost-benefit analysis are the benefits and potential issues of using VGI.

Benefits of using VGI for CGDI include:

- Cheap to no cost GI at a very high speed of updates;
- Source of up-to-date geographic information;
- Huge number of possible contributors;
- Recent studies showing that VGI is capable of meeting positional accuracy standards of authoritative datasets under certain conditions;
- Opportunity to produce richer and more up-to-date databases.

Potential issues of using VGI for CGDI:

- Most VGI is not accompanied by a metadata and can be of poor quality. Findings of this research, discussed in Chapter 4, demonstrate that while it is possible to verify the positional accuracy of VGI, the problem of validating attribute accuracy and metadata still remains. Subsequently being incapable of validating all aspects of VGI quality and using it to update authoritative datasets could lead to a situation where a user could be harmed by relying on such information.

However the general reliability of VGI data and/or the ability to obtain large quantities of data at a low cost means that it's not unreasonable to incorporate VGI to update CGDI datasets. For example the Province of Saskatchewan has completely stopped updating their topographic database while the Province of New Brunswick hasn't updated their topographic database since 1998 mainly due to high cost of maintenance. Possible use of VGI to update such authoritative datasets could be seen beneficial.

A report entitled "Examination and Critical Comparison of Alternative Maintenance Models for the Nova Scotia Digital Topographic Database" [Tienaah et al., 2013], summarized work in the field of collection and maintenance of authoritative datasets undertaken by ten Canadian government mapping organizations and three international mapping organizations. It showed that all Canadian provinces are investigating into new technologies and services for data acquisition (including possibility of using VGI) and that 80% of examined mapping organizations maintain their topographic databases by relying on data coming from their partners, where the quality of such data is later reviewed by an authoritative mapping organization. In the report, the move by authoritative mapping organizations from collection of GI on their own to utilization of other existing sources of GI and their validation for maintenance of authoritative datasets has been identified. The main justifications for implementing such practices were due to high cost of datasets maintenance and substantial budget cuts [Tienaah et al., 2013].

But does it mean that because of budget cuts and as a result substantial decrease in updating of authoritative datasets it would be reasonable for authoritative mapping organizations, like CGDI, to use VGI as a source of updates? To answer this question, it is necessary to analyze standard of care element when integrating VGI. The application of standard of care in this situation could be influenced, but not determined, by:

> *proof that the defendant's behavior was consistent with the established practices and customs of other citizens carrying our similar activities and tasks…. The current approach is to view proof of compliance with approved practice as providing some evidence of due care. The strength of that evidence depends upon the longevity of the practice, its universality, the status and reputation of the profession or occupation and its members, the degree of technical difficulty of the task at issue, and any evidence of additional precautions that may have been available.*
>
> [Osborne, 2007]

The Nova Scotia report [2013] provided a summary of technology that was used by examined authoritative mapping organizations in maintenance of their topographic databases, and is shown in Figure 2.3.

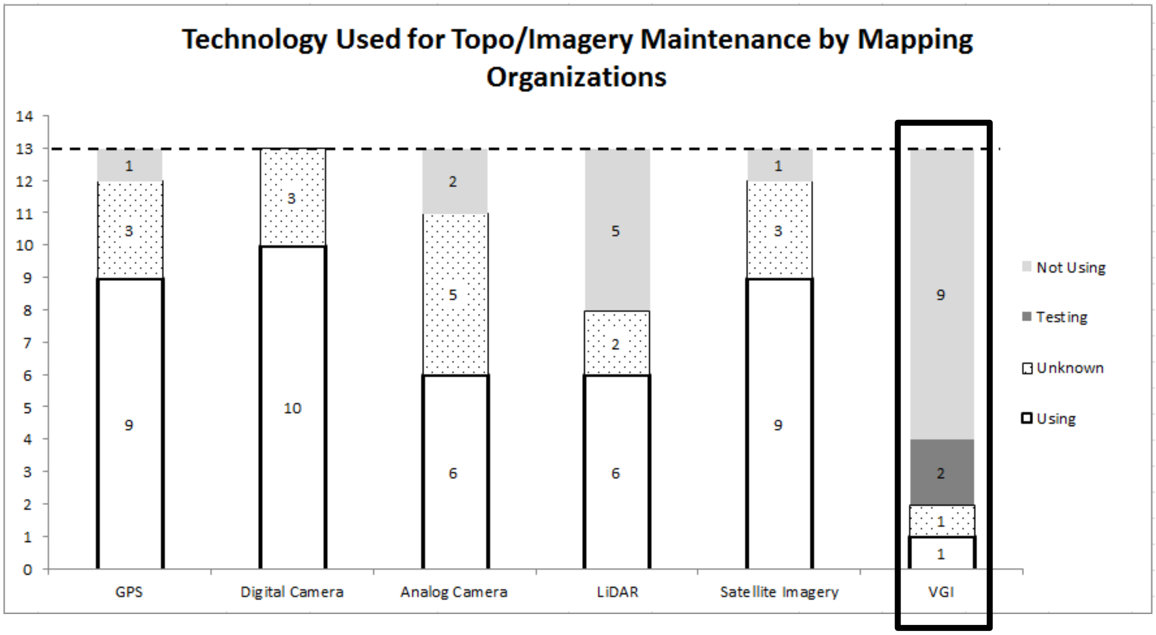


Figure 2.3 Tehcnology Used in Topographic and Imagery and Maintenance (from [Tienaah et al., 2013])

In Figure 2.3, the horizontal axis represents technologies and the vertical axis represents number of organizations examined during the study. Figure 2.3 illustrates that out of thirteen mapping organizations examined, none of the ten Canadian mapping organizations utilized VGI as a source of updates to their datasets, and only two of them are testing the possibility of such integration. Out of three international mapping organizations examined only one (Notification for Editing Service, State of Victoria, Australia) actually uses VGI, and only for the purposes of notification and not direct update of the datasets. The report [Tienaah et al., 2013] shows that in Canada the use of VGI as a source of updates to authoritative datasets is still not a common practice, although some of the mapping organization showed their interest in using VGI and have started initial testing. The current approach to such customary practice is to "view proof of compliance with approved practice as providing some evidence of due care" [Osborne, 2007].

> *A deviation from customary standards is not conclusive of negligence but it is likely to be given significant weight by the court, and the defendant is wise to negate this evidence by giving some explanation for the deviation or by showing that her conduct was no more dangerous than the standard practice.*
> [Osborne, 2007]

In the light of such circumstances the decision of a judge, to treat the possible decision of CGDI to update their datasets using VGI as an unreasonable one (breach of standard of care), would not be highly unforeseen. However, if the court finds that the customary practice itself is unreasonable, a "finding of negligence may be made" [Osborne, 2007]. Decision, on whether use of VGI as a source of updates for CGDI datasets is reasonable (no liability) or unreasonable (liability), will also be influenced by specific circumstances of a case based on foreseeability of risk, likelihood and seriousness of damage, cost of preventive measures, utility of defendants conduct, judicial policy, economic analysis and others [Osborne, 2007].

The last three elements of negligence (causation, remoteness and damage) would need to be established in a case when physical harm has been caused to a user by reliance on VGI used as a source of updates for CGDI datasets. The plaintiff will have to prove that the negligence of the defendant caused his harm (causation element). The conventional test to determine the causation element is the *but for* test where "one must ask the question "would the plaintiff's damage have occurred *but for* the defendant's negligence" [Osborne, 2007]. If the answer is "yes" than the causation element of defendant's negligence would be established. The next negligence element, damage, requires that an actual harm has been "suffered by the plaintiff as a result of defendant's wrongful act… negligent conduct without consequences will not do" [Osborne, 2007]. The plaintiff will also have to prove that harm was not too remote from the negligent act of defendant (remoteness of damage element), and will depend on whether the injury was a foreseeable consequence of a breach of the standard of care. The decision on this element of negligence will most likely to be based on a "current judicial policy and intuitive and impressionistic sense of fairness about where to draw the line on the defendant's responsibility" [Osborne, 2007]. The three above mentioned elements of negligence would be quite straightforward to prove.

Once the plaintiff has provided evidence and a judge or a jury agreed that elements of negligence took place, the defendant (CGDI) can be held directly liable in negligence to users who are injured due to errors introduced by VGI. But it should be noted that "whether the defendant was negligent depends on the circumstances of the particular case, and the fact that defendants have been found liable, or not liable, in previous cases which seem similar, does not mean that a similar result will necessarily follow in subsequent case" [ICBCclaiminfo, 2013].

Liabilities in tort for VGI most likely arise in organizations that incorporate VGI into their datasets (such as CGDI) and their users, and not the volunteers of VGI. "Defendants in the tort must be worth suing" and as mentioned by Osborne [2007] such defendants usually fall into one of the following three categories:

- People or institutions with liability insurance (insured defendants);
- Large corporations or governmental institutions (self-insured defendants);
- Uninsured persons or institutions that have sufficient assets (defendants with means).

Contributors of VGI do not fall within any of these categories, and would not be considered as financially viable defendants. Also Scassa [2013] encompasses that VGI is mostly combined from thousands of "disparate sources to form one or more coherent work" and it would be practically impossible to find an individual to be signed out for liability. CGDI, led by the Natural Resources of Canada, on the other hand can be considered as a "financially viable defendant" [Osborne, 2007].

2.7 Controlling Liability and Risk Management Techniques

After analyzing the literature review on legal liability, a definition of legal liability in VGI for this research is proposed. Legal liability represents the responsibilities of an organization arising from contract or from civil actions (torts), as a result of the distribution, or other forms of dissemination of VGI, in a case where such errors caused physical harm, and the user made a successful claim to a court. To avoid and limit potential liability for CGDI from the integration of VGI into the updating process of CGDI datasets, a set of risk management techniques is required. "Risk management is any procedure or practice that … organizations use to reduce exposure to liability" [Bertrand and Brown, 2006]. To avoid liability, an information provider must be "able to show that reasonable steps are routinely taken to ensure the accuracy of disseminated information" [Low et al., 2010].

Different organizations manage their liability risks differently [McCurley and Lynch, 1996; Graff, 2003; Martinez, 2003]. Each organization is complex and unique, so "one-size-fits-all risk management programs are not possible" [Bertrand and Brown, 2006]. This section describes four risk management techniques to reduce/limit legal liability risks around VGI, which has been described in Sections 2.5 and 2.6. The following risk management techniques are summarised in Table 2.1.

Table 2.1 VGI Risk Management Techniques

No.	Name of Risk Management Technique	Description
1	Identification of Possible Risks	- Form a risk management team - Identify and rank appreciable risks - Evaluate likelihood of risk occurrence against magnitude of harm - Develop processes/stages to minimize and reduce identified risks - Apply them to identified risks - More regular checks on samples of data, to verify if risks are still there
2	Quality Assurance Procedures	- Overlay with other spatial datasets - Use satellite / aerial imagery - Use rating / feedback system - Manage datasets containing VGI separately
3	Disclaimers in Contract	- Use liability disclaimer - Consult with experts
4	Duty to Warn about Quality	- Warn about VGI quality before use - Warn about VGI quality after use

Additional information on risk management techniques can be found in the work of Marc Gervais [2012] who conducted an extensive Pan-Canadian survey about data quality and various possibilities of managing risks in geomatics. Also the use of risk-management approaches to minimize the risks of inappropriate usages of data by non-specialists has been conducted by Bédard et al. [2013], and different means, like certification, accreditation, inspection, audit, warranty, quality assessment, quality control, quality assurance, to reduce risks of geospatial data misuses which can be applied to VGI by Larrivée et al. [2011].

2.7.1 Identification of Possible Risks

If an organization develops an RMP that identifies only a few risks, there is a chance that more important issues will be omitted due to the fact that not enough time has been devoted to this step [Tremper and Kostin, 1993; Safritt et al., 1995]. On the contrary, if "a list includes every conceivable risk imaginable [it] is almost as useless as having no risk at all" [Martinez, 2003]. GI providers should set up an applicable risk management process to reduce the liability concerns surrounding implementation of VGI, comprising of the following steps:

- Develop a risk management team [Martinez, 2003];
- Identify and rank appreciable risks [Martinez, 2003];
- Evaluate likelihood of risk occurrence against the magnitude of harm [Martinez, 2003];
- Develop processes/stages to minimize/reduce the most important risks;
- Apply them to identified risks;
- Perform regular checks on samples of data, to verify if risks are still there; and
- Continue identification of new risks.

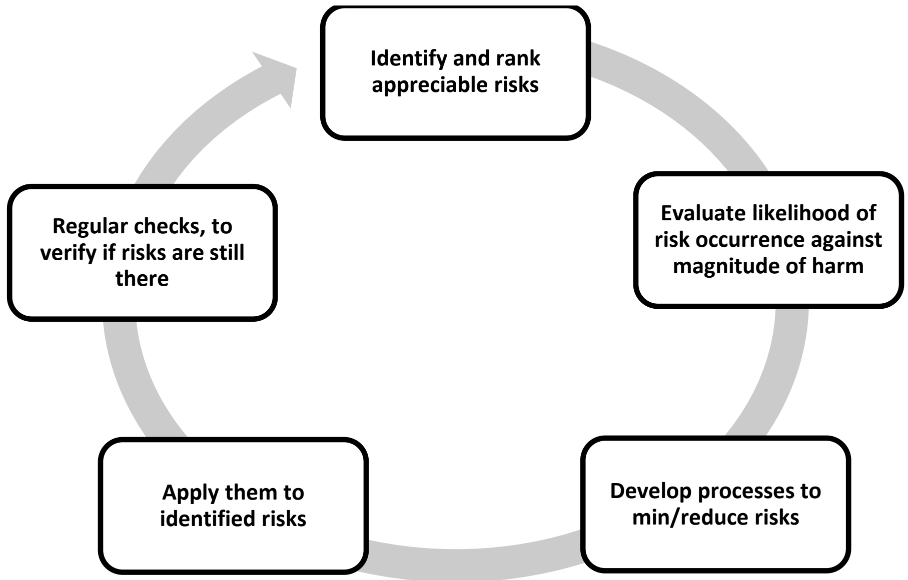


Figure 2.4 Identification of Possible Risks

2.7.2 Quality Assurance Procedures

Organization is required to keep a constant focus on GI quality, including incorporated VGI. The GI organization can and should employ the newest available techniques to check the quality of contributed GI such as checking VGI quality against trusted datasets and/or their own, overlaying VGI data on top of aerial/satellite imagery, or using a rating/feedback system that enables expert users in the community to provide feedback concerning the contributed VGI. The GI providers should devote their time and resources to check for VGI quality. Another important risk management technique for CGDI could be maintaining a number of GI datasets, such as:

- A dataset that is being updated using traditional methods. This dataset should be less prone to errors, and legal liability in this case will not likely arise.
- A dataset that is being updated using VGI, and GI from traditional methods when those become available. Such a dataset will have a faster rate of updates, but will have a higher chance of errors in it.
- A dataset that is being updated using VGI only. Such dataset has the greatest potential to contain errors, and as a result imposes higher risks of legal liability on CGDI. In this case it will be very important for CGDI to communicate to its user a message of quality of such datasets, and explain that a choice of dataset should be based on their own "purpose for use".

Such risk management techniques would limit/reduce the potential for legal liability of CGDI, as it will be the user who will make an informed decision which dataset to use.

2.7.3 Usage of Disclaimer in the Contract

GI providers cannot completely limit and restrict their liability in the contract from all possible damages the use of VGI can bring [Chandler and Levitt, 2011]. With the view that VGI can be used in various and unpredictable ways to serve different purposes for users, therefore the GI providers can indicate in the contract or liability disclaimer that before using VGI the user should consult with an expert and/or GI provider whether such data can be used for the intended purpose [Chandler and Levitt, 2011].

2.7.4 Inclusion of Duty to Warn about Quality of VGI

As pointed out by Chandler and Levitt [2011] and Gervais [2003], at a minimum the GI provider should always exercise a duty to warn the user/receiver of the information about the quality of the data and indicate any limitations that might come with it. The liability disclaimer should clearly point out that the volunteered information is to be used only for certain purposes, or be employed at the user's own risk [Club TomTom, 2007]. It is preferable to indicate disclaimers not only when data are downloaded from the website or other source by the user, but also to constantly have a communication process between GI provider and user [Chandler and Levitt, 2011]. The organization should keep contact details of their users, at a minimum an e-mail address, to be able to warn users of any new risks or known errors concerned with the use of VGI, which might not have been discovered.

The above mentioned risk management techniques are directed towards preventing CGDI from exposure to legal liability. When a case of legal liability is inevitable a purchase of insurance is recommended. Few agencies have the "reserves or funds necessary for complete self-insurance of their exposures" [Bertrand and Brown, 2006].

Although it should be noted that not all risks are insurable and CGDI should study all possibilities before purchasing one.

2.8 Summary

The background information provided in this chapter is intended to place this research project within a larger context. Description of data quality concepts and their implications on VGI were given, and were followed by a review of the current validation techniques for assessing quality of both authoritative GI and VGI. The chapter went on to show that incorporation of VGI within authoritative datasets might create new legal risks for CGDI, with legal liability at the forefront. The chapter provided a review of the required characteristics to which VGI should comply in order to be integrated with CGDI.

The chapter concluded with a preliminary analysis of the legal liability risks for CGDI in contract and tort, as well as an analysis of risk management techniques that can be incorporated by CGDI to minimize/reduce such risks.

3. Conceptual Design and Implementation

The first chapter identified the need for integration of VGI with authoritative datasets, like the ones maintained by CGDI, so more up-to-date, cost-efficient GI datasets could become available. Such GI datasets could be achieved by utilizing more VGI and fewer traditional data. A prerequisite for such integration of VGI and CGDI datasets is identification and selection of VGI that corresponds to the CGDI accuracy standards. The integration process of VGI and CGDI, as described in Chapter 2, is also dependent on the legal issues that might arise from such integration, with legal liability at the forefront.

The goal of this chapter is to describe the methods that were used in this research to meet the thesis objectives on design and deployment of a prototype (VGI_Editor) and implementation of Risk Management Techniques (RMT) that could reduce/limit potential legal liability concerns surrounding integration of VGI and CGDI. This chapter discusses the conceptual design of the prototype that facilitates integration of VGI and CGDI datasets. It then describes design execution of the VGI_Editor prototype using publicly available web mapping tools. The chapter concludes with a discussion on the spatial data format used within the VGI_Editor application.

3.1 VGI_Editor

To achieve the research goal the prototype designed in this research should be capable of performing the following functions using generic needs from published research on integration of VGI with authoritative datasets:

- The prototype should allow uploading VGI in vector format, including points, lines and polygons;
- The prototype should allow for visualization of VGI;
- The prototype should provide digitizing and editing function for the creation of spatial content;
- The prototype should provide a means for making comments and ratings of VGI contributions and their contributors by other users;
- The prototype should provide a means of validating user contributions;
- The prototype should provide a means for extracting VGI from the database;
- The prototype should be providing a set of risk management techniques to limit/reduce potential legal liability issues around VGI;
- If time permits - the prototype should provide buffer, intersect and statistical analysis of VGI with data from authoritative datasets.

The above mentioned functions of the VGI_Editor prototype were achieved in full, except for the last function - the prototype should provide buffer, intersect and statistical analysis of VGI data, which have been implemented using ESRI ArcGIS software - and are described in the following sections of Chapter 3.

Design of the VGI_Editor prototype was not chosen for “its good looks; instead, it concentrates on getting the features working” [Westra, 2010]. One of the main characteristics of the VGI_Editor prototype is the validation of VGI for compliance with CGDI accuracy standards. Another area of focus in development of the prototype is incorporation of RMTs to handle legal liability issues surrounding integration of VGI with CGDI datasets. Figure 3.1 represents the role of VGI_Editor in the process of VGI validation to meet the CGDI accuracy standards.

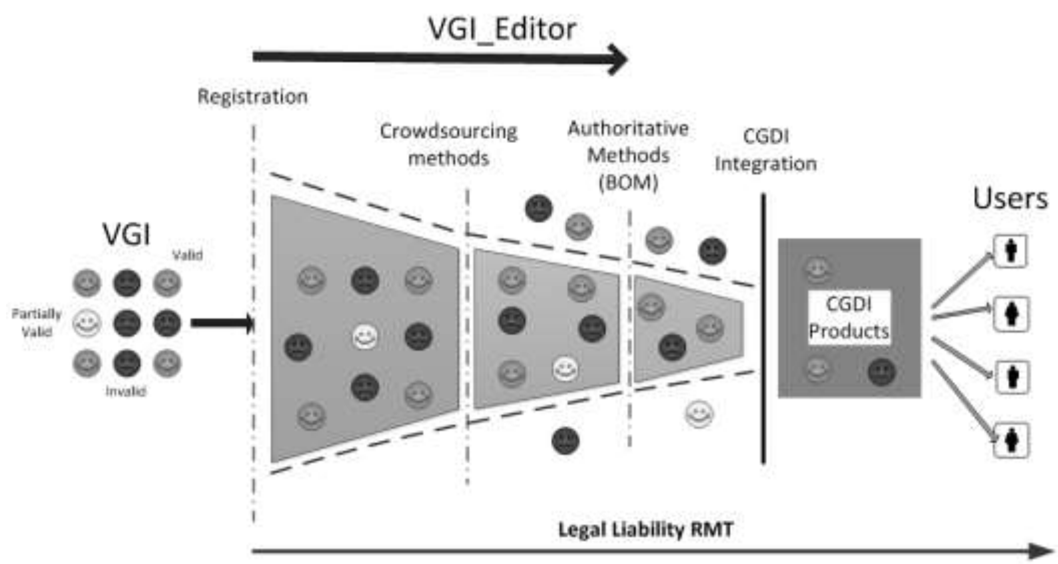


Figure 3.1 Role of VGI_Editor in the Process of VGI Evaluation to Meet the CGDI Accuracy Standards (after Overton and Abbott [2007])

The validation of VGI is proposed to start at the point when the contributor is willing to submit collected VGI into the VGI_Editor. VGI, inside VGI_Editor, will go through crowdsourcing and authoritative methods of assessing the quality of GI. In the case of satisfying data quality requirements, such VGI could potentially be used to update CGDI datasets. In this thesis, the author argues that the combination of an authoritative approach such as the Buffer Overlay Method (BOM) used to assess the quality of GI with crowdsourcing methods, such as ratings, peer reviews and other methods used by wiki projects, would be enough to meet the accuracy standards for VGI. If this hypothesis is proven wrong more research into validation methods of VGI quality would be required. The combination of two approaches includes:

- Crowdsourcing methods
    - Crowdsourcing work on geometry and attributes of VGI,
    - Voting system,
    - Comments, and
- Authoritative method
    - BOM.

The dotted black line in Figure 3.1 represents the boundary of the "data quality step". The amount of data is largest at the first stage of the VGI_Editor because of an "anything goes in" approach of the VGI_Editor prototype. The only requirement for submitting VGI is the registration of the user within the VGI_Editor prototype. Research led by Nkhwanana [2009] suggests that a registration system should be established when validating VGI. Such a registration system should not compromise the contributors' privacy and/or security, but should rather allow "users to be identified only for the purposes of reliably identifying trustworthy contributions". Some of the problems that might arise during this stage are the problems of verifying identity of contributors who use "meaningless nicknames". Some of steps that could prevent this from happening, that are not a part of this research, include recording of IP address of the contributor, and/or requiring, for example, LinkedIn or Facebook login credentials of the contributor. At the next stage, as shown in Figure 3.1, VGI would go through crowdsourcing methods of validating VGI, followed by an authoritative BOM, which would potentially filter out VGI that does not satisfy the CGDI accuracy requirements.

A solid black line in Figure 3.1 represents the point at which the VGI data has been integrated with the data from the authoritative datasets. As noted by Nkhwanana [2009] "automatic methods are less likely to be adopted for use in authoritative databases, where quality and credibility are of utmost concern. The final verification of contributions in authoritative databases has to be carried out internally by the administrators". Once administrators of CGDI datasets have made a decision on whether VGI satisfies the accuracy standards, e.g. position, shape, attributes, categorization of the object, it could be used to update datasets maintained by CGDI. The process of actual integration of VGI and CGDI datasets is not covered in this research. Before distributing CGDI datasets updated with VGI, administrators of such datasets should make sure that RMTs are in place to limit their legal liability, as VGI might introduce errors to CGDI datasets.

3.1.1 VGI_Editor - Quality Compliance

As mentioned in Section 3.1, one of the main characteristics of the VGI_Editor is the validation of VGI to correspond to CGDI accuracy standards. Figure 3.2 represents the stages VGI goes through after submission to the VGI_Editor in order to comply with CGDI accuracy standards.

At the first stage users register through the WebGIS application. Findings of Nkhwanana's [2009] research suggest that a registration system requiring an email address is preferred. It provides a means of communication, as well as the ability to discover and track future contributors.

After the registration user is given an option to upload VGI into the VGI_Editor. A volunteer can upload data in an ESRI shapefile format, which is GIS most popular geospatial vector data format. At this step, users will have an option to fill in a metadata file for uploaded VGI. Metadata can include basic information such as source, time and any other information deemed necessary by the user, and will be used in a form of tags, as described in more details in Section 3.4.5.

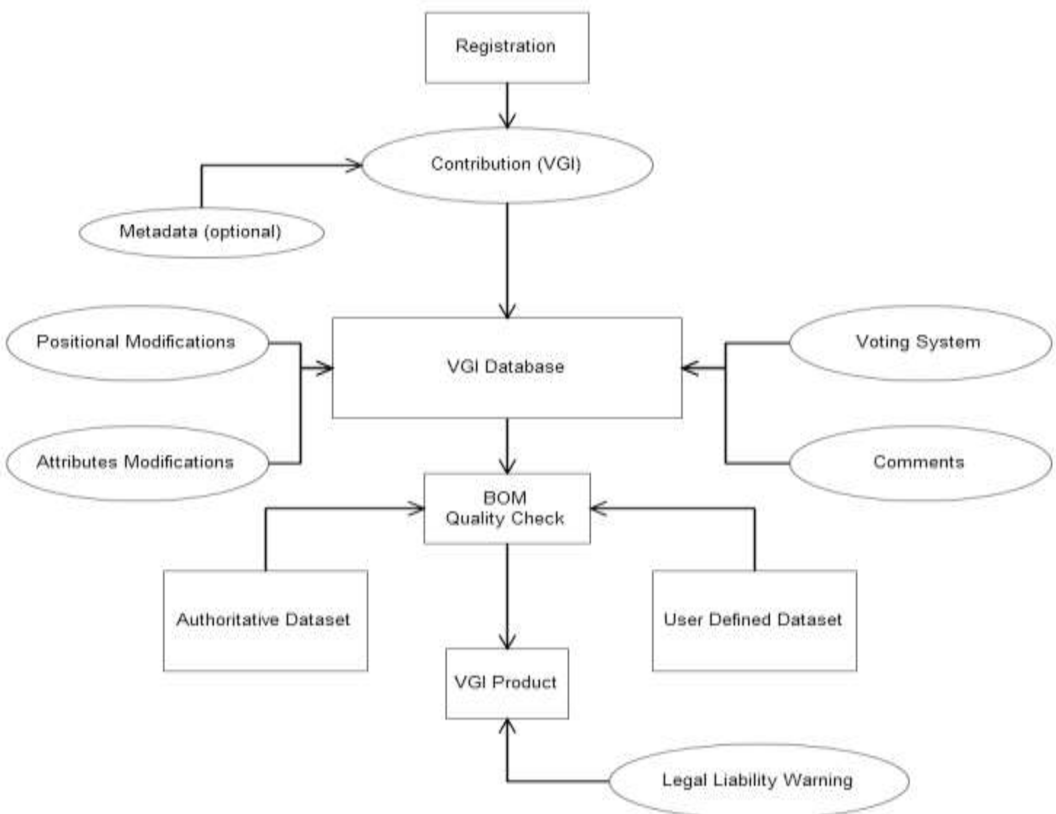


Figure 3.2 VGI_Editor Design

In this step, data successfully uploaded to the VGI_Editor database is reviewed by the registered users of WebGIS prototype. Users are given options to modify the geometry and attributes of vector objects (add/delete/modify), as well as to leave comments. VGI could also be rated using a simple like/dislike voting system that is described in more detail in Subsection 3.4.6.1. A possible future update to the prototype could also include the finding of Jones [2011] on using 5-star rating system on crowdsourced geospatial data quality assessment. It could provide more information on the VGI (5 values instead of 2), including such information as an average, the distribution of votes per page, the most useful, best and worst comments, etc.

Verified VGI can be later downloaded from the VGI_Editor portal. At this stage the potential users are responsible for ensuring that VGI satisfies their needs, by using tools provided through the VGI_Editor. After ensuring that VGI satisfies their needs, the data can be downloaded. Novice users seem to assume spatial data is accurate, and thus must be notified about the existence and significance of uncertainty and error [Chandler and Levitt, 2011]. Given that the duty to warn is a “continuing obligation”, information about spatial data quality should accompany spatial datasets [Gervais, 2003; Chandler and Levitt, 2011]. The risks associated with using VGI will be provided within the VGI_Editor prototype to address the legal liability concerns.

3.1.2 VGI_Editor - Legal Liability Risk Management Techniques

Another main characteristic of the VGI_Editor is the incorporation of RMTs to handle legal liability issues surrounding implementation of VGI with CGDI datasets. In order for CGDI to limit and/or reduce their exposure to legal liability, because of VGI integration, CGDI should be "able to show that reasonable steps are routinely taken to ensure the accuracy of disseminated information" [Low et al., 2010]. One RMT of primary importance is the use of the VGI_Editor as a linking tool in the integration process between VGI and CGDI. The VGI_Editor can be looked at as a "coherent protocol in place to ensure the accuracy of both the agency supplied source information and information products that are an aggregation of multiple source data sets" [Low et al., 2010].

The VGI_Editor, as described in section 3.1 and subsection 3.1.1, provides tools for checking the quality of VGI and as such helps users to make an informative decision on GI data quality. A more detailed utilization of RMTs used by the VGI_Editor will be described in Subsection 3.4.7.

3.2 VGI_Editor Architecture

VGI_Editor adopted principles, guidelines and code described by Westra [2010] in his book "Python Geospatial Development" development of *ShapeEditor* prototype*;* and

Django Bookmarks prototype described by Hourieh [2009]. The VGI_Editor has been developed as a Web GIS system, which is an "…implementation of geographic information functionality through a web browser or other client programs, thus allowing a broader usage and analysis of a particular geographic database" [Tienaah, 2011], that follows rules and standards of Open Geospatial Consortium (OGC), except for data exchange format which will described in Section 3.5. Development of the VGI_Editor is based on OGC standards because they "… support interoperable solutions that 'geo-enable' the Web" and also on the CGDI itself that was an initiative of OGC's Interoperability Program [Open Geospatial Consortium, 2012]. As a result, OGC standards will facilitate not only the processes inside the VGI_Editor, but could also ensure better interoperability between the VGI_Editor GI and CGDI datasets. The VGI_Editor for the purposes of this research was developed by using free and open source GIS Web tools. The VGI_Editor prototype is developed as an evolutive prototype that would become the final system after a number of modifications to the system based on future implementation of user requirements.

VGI_Editor involves three application tiers:

- Client tier (deals with user interface, serves as the output and input interface to users) [Westra, 2010; Tienaah, 2011];
- Application tier (the middle tier for communication between the client and data tier); and
- Data tier (deals with the storage of data).

Subsections 3.2.1, 3.2.2 and 3.2.3 give a detailed description of each tier.

3.2.1 Client Architecture

The Client tier serves as an interactive interface for the VGI_Editor users. Through the Client tier, users of the service can interact with GI and other services that are provided by the VGI_Editor. As noted by Tienaah [2011], when choosing the right tool for a Client tier, one should keep in mind the following characteristics of the tool:

- Application Programming Interface (API),
- Community support,
- Data, or support services,
- Standards (interoperability),
- Functionality,
- Flexibility, and
- Licensing agreement.

Current trends indicate an increased adoption of OpenLayers as a base platform for developing web-mapping platforms [Adams, 2009; Tienaah, 2011].

> *Another key advantage of OpenLayers is that, it provides a means of imbedding publicly available map clients such as Google, Bing and Yahoo maps. These map clients provide a wide public familiarity to map navigation and usage.*
>
> [Tienaah, 2011]

Figure 3.3 represent the connection between above described tier and system components.

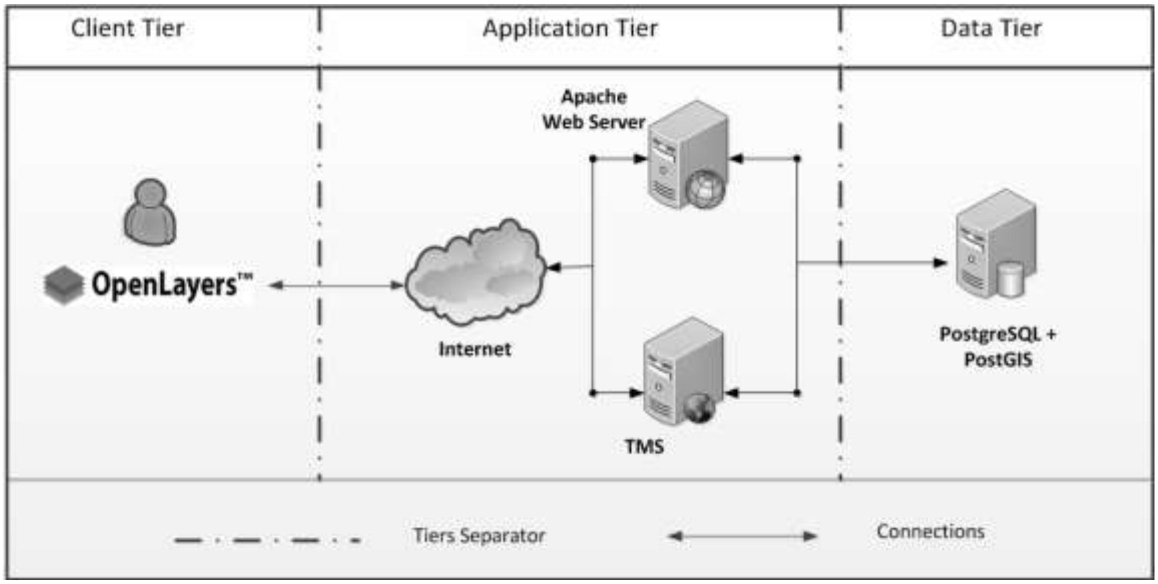


Figure 3.3 VGI_Editor System Tiers (after [Tienaah, 2011, p. 49])

3.2.2 Spatial/Web Server Framework

The Application tier of the VGI_Editor represents the "business logic" of the application; it serves as a link between requests coming from the Client tier, and the corresponding data stored inside the Data tier. The Application tier plays an important role in the VGI_Editor, as it allows users to create, read, update, and delete records stored inside the Data tier. The VGI_Editor utilizes the following frameworks as components of the Application tier:

- Apache – web server framework, and
- Tile Map Service (TMS) - spatial server framework.

The web application server “listens for incoming HTTP requests from web browsers and returns either static content or the dynamic output of a program in response to these requests” [Westra, 2010]. An *Apache* [Apache, 2012] web server has been used with the VGI_Editor, as it is one of the most popular free web servers available. The VGI_Editor prototype addresses both spatial and non-spatial data. A spatial server, used to generate and serve spatial content from Data tier to Client tier, is required for a thoroughly functional application [Westra, 2010]. The VGI_Editor utilizes a TMS protocol, which “defines the interface to a web service that returns map tile images upon request” [Westra, 2010]. The VGI_Editor is based on a RESTful (Representational State transfer) web service protocol, which “uses parts of the URL itself to tell the web service what to do” [Westra, 2010]. A Uniform Resource Locator (URL) is the “global address of documents and other resources on the World Wide Web” [Webopedia, 2012]. By using RESTful principles, the URL used to access the web service contains all necessary information for rendering a tile and returning the requested map back to the user.

3.2.3 Data Tier

“Relational database management applications dominate the field of data storage and retrieval” [Tienaah, 2011]. In this research an open source object-relational database system *PostgreSQL* [PostgreSQL, 2011] with a *PostGIS* extension has been chosen as a

Data tier for VGI_Editor prototype. This particular extension “spatially enables” the *PostgreSQL* server, allowing it to be used as a backend spatial database.

PostgreSQL database with *PostGIS* extension has been chosen for the following reasons [PostgreSQL, 2012]:

- Storing and manipulation of spatial objects like any other object in the database;
- No licensing cost for the software;
- Reliability and stability;
- Cross platform;
- Designed for high volume environments; and
- Wide user support and functionality.

The VGI_Editor database model is made of twelve inter-related tables. Figure 3.4 represents a Unified Model Language (UML) diagram showing the relationship between tables inside the database.

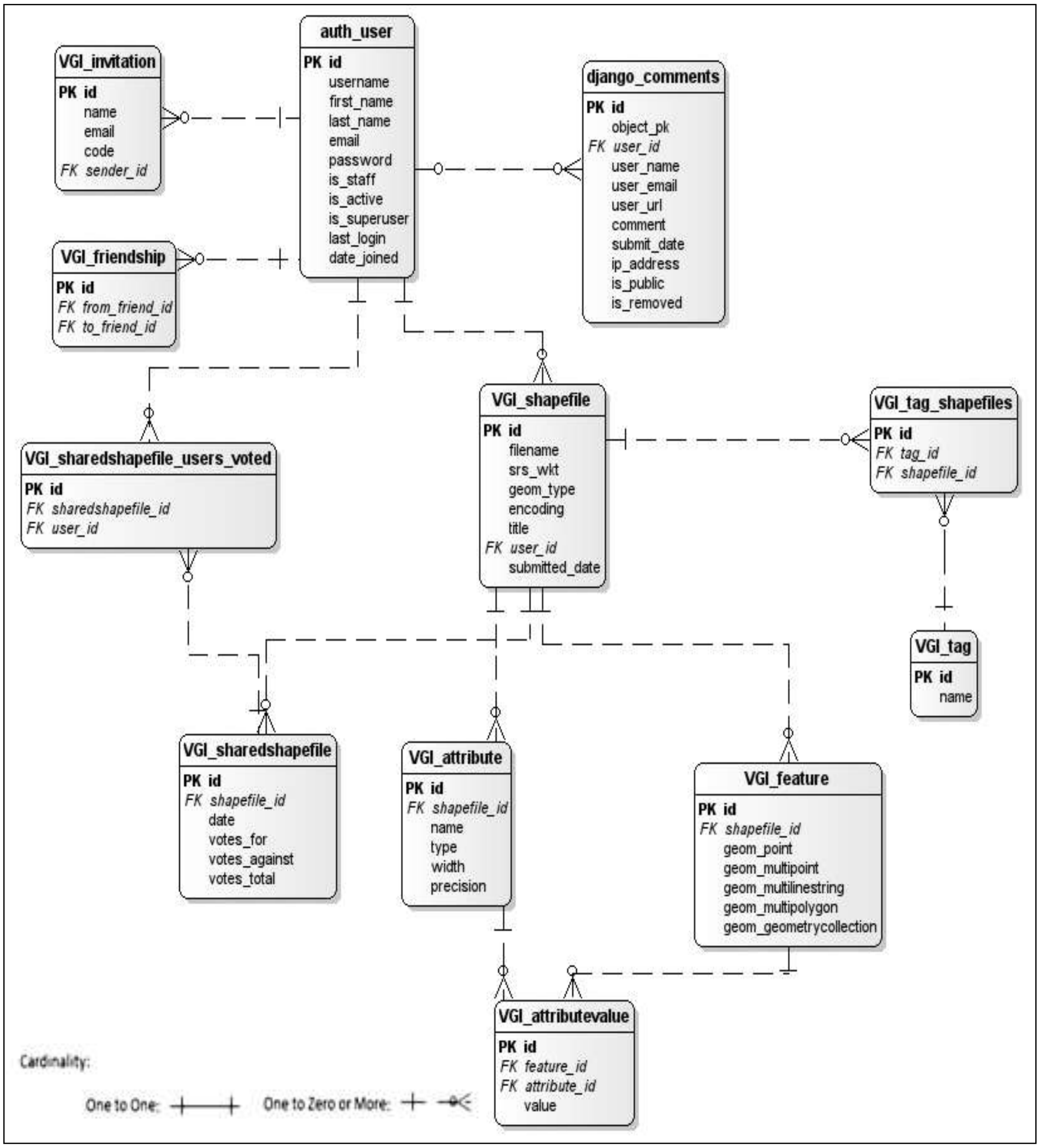


Figure 3.4 VGI_Editor Database Model Diagram

3.3 Choice of Web Application Framework

A Web Application Framework (WAF) is a "software framework that is designed to support the development of dynamic websites, Web applications and Web services" [Wikipedia, 2012]. WAFs often save a lot of time when building web sites, as they provide "core functionality common to most web applications, such as user session management, data persistence, and templating systems" [DocForge, 2012].

According to DocForge [2012], some of noteworthy WAFs include the following:

- For Java
    - Apache Struts
    - JavaServer Faces
    - Jt Design Pattern Framework
    - Apache Wicket
- For PHP
    - CakePHP
    - CodeIgniter
    - Symfony
    - Zend Framework
- For Python
    - Django
    - Flask
    - Pyjams

    - Web2py
    - Pylons
    - Turbogears
    - Twisted
    - Web.py
    - Zope
    - Pyroxide
- For Ruby
    - Ruby on Rails
    - Ramaze
- Other
    - .NET

The VGI_Editor has been developed using Django, a Python-based WAF that supports geo-spatial web application development. GeoDjango is a set of extensions to Django that enables geo-spatial capabilities of the framework. The GeoDjango WAF was chosen over others for the following reasons [Hourieh, 2009; Westra, 2010]:

- It is based on the Python scripting language, providing cross-platform development with great readable syntax.
- There is tight integration between components.
- It includes an Object-Relational Mapper providing a bridge between the data model and the database engine.
- It has a clean URL design (are both user and search engine friendly).

- It provides an advanced development environment (comes with a lightweight web server for development and testing).
- It is a mature WAF with a great documentation and release process.
- It is very stable, with backwards-compatibility policy.
- It has a great number of libraries (“batteries included”), including: user authentication, session management, syndication, templates, administration tools, forms, caching, testing, etc.

A Django project is based on a number of applications, where each application consists of three main parts [Westra, 2010]:

- Model (contains everything related to the application’s data, including structure, import, access, validation and son on);
- Template (describes how information will be presented to the user);
- View (responsible for business logic of the application, usually utilizes model and template to produce the output).

Django’s model-template-view architecture, as shown in Figure 3.5, loosely follows the model-view-controller (MVC), as shown in Figure 3.6. This is a widely-used pattern in software development which keeps “the application component of the software independent from the interface, and a general mechanism binds the user interface to the data model without introducing interface knowledge into the data model” [Tienaah, 2011]. The benefits of this pattern are as follows [Hourieh, 2009]:

- Designers can work on the interface without worrying about data storage or management.

- Developers are able to program the logic of data handling without getting into the details of presentation.

> *The two (Django's model-template-view and MVC) are quite distinct, and describe the different tiers in the web application stack in very different ways. While the model in both Django and MVC represents the data tier, Django uses the view to hold the application logic, and separates out the presentation using templates. MVC, on the other hand, allows the view to directly specify the presentation of the data, and uses a controller to represent the application's business logic.*
>
> [Westra, 2010]

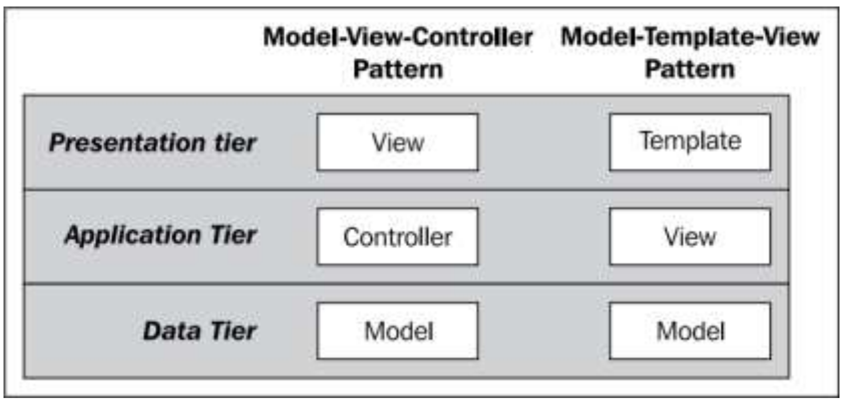


Figure 3.5 Django's Model-Template-View Pattern vs. MVC (from [Westra, 2010])

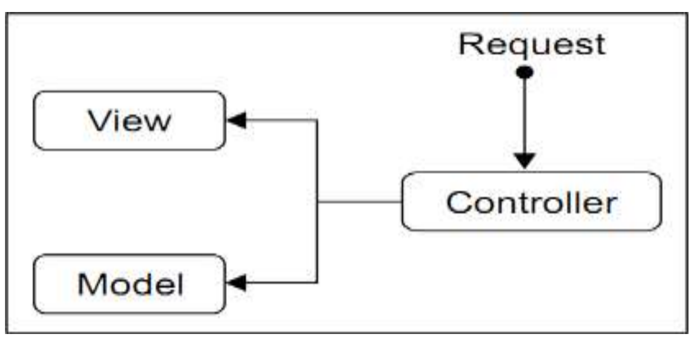


Figure 3.6 MVC Pattern Interactions with User Request (from [Hourieh, 2009])

3.4 Client Side Implementation

The Client side implementation is based on an OpenLayers Web GIS platform, a JavaScript library for displaying maps on the client side. The client interface implementation consists of eight visible views, as shown in the Figure 3.7. The function of each view is described in sections 3.4.1 – 3.4.7

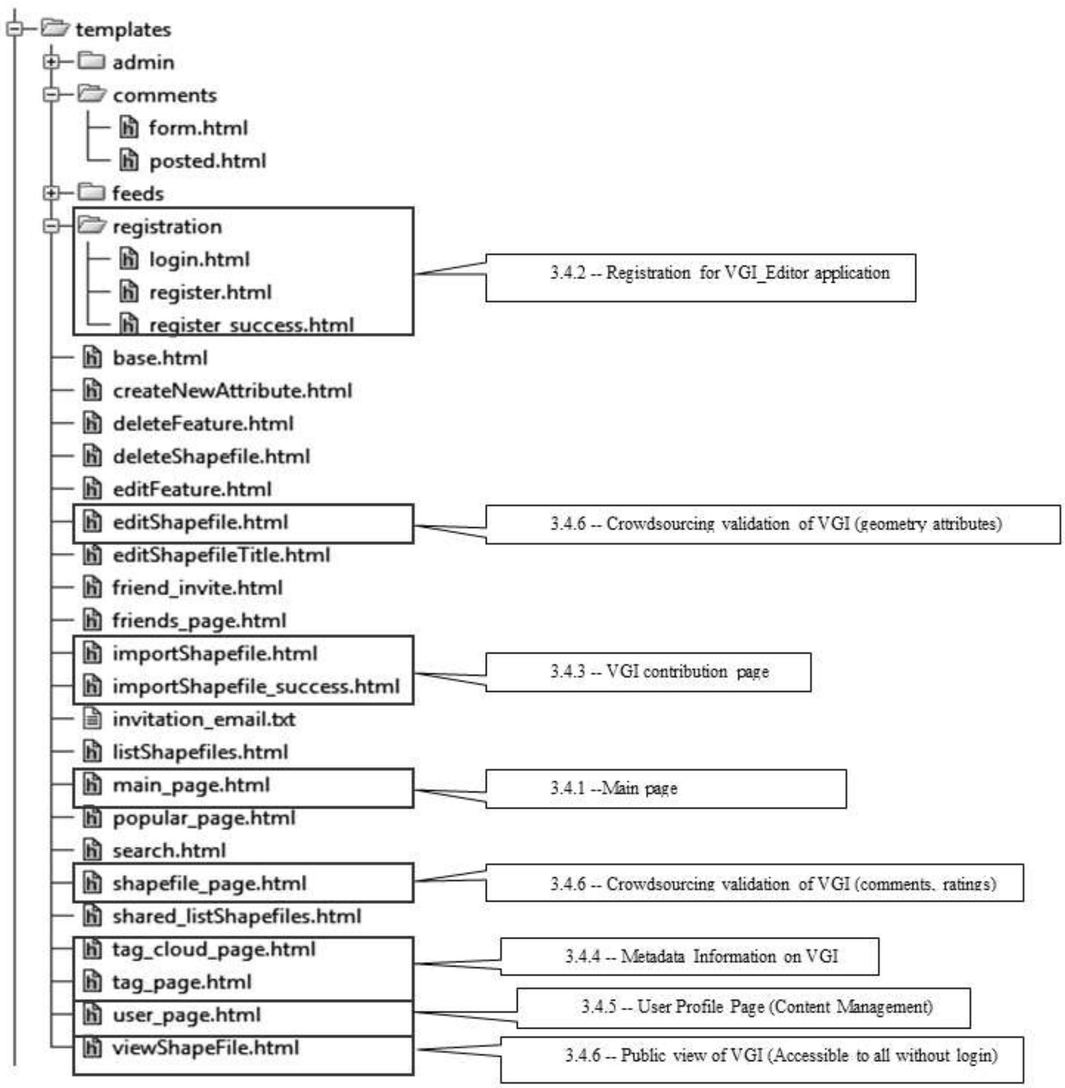


Figure 3.7 VGI_Editor Views

3.4.1 Main View

The Main View page serves as an entry point to the VGI_Editor prototype, and its basic layout for registered users is as shown in Figures 3.8 and 3.9. The purpose of this page is to serve as the interface for all users who have entered the website, including both registered and non-registered users. Main View for the non-registered users is shown in Figure 3.8, and gives access to the following functions of the prototype: registration, login, tag cloud (metadata), recent (the most recent submitted VGI), and list of VGIs that have been shared with users of VGI_Editor prototype (discussed in more details in Subsection 3.4.5).

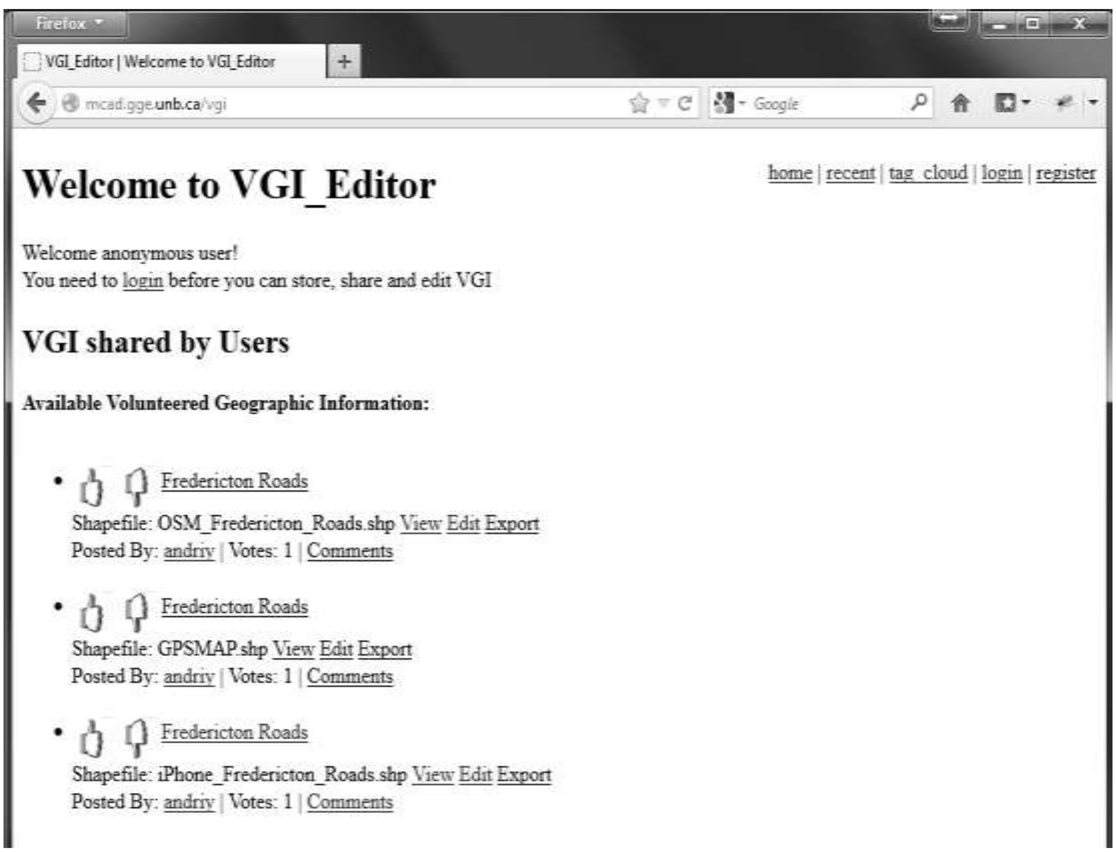


Figure 3.8 Main View for Non-Registered Users

Registered users can login and can proceed to the following functions of the prototype: logout, search, user profile and contribution pages. The Main View page layout for registered users is shown in Figure 3.9.

Figure 3.9 Main View for Registered Users

3.4.2 Registration View

"The process of registration is one in which there is a recording of the relevant data about a person in a specific matter of importance. In other terms this can be also referred to as a process of enrollment or even signing up for a purpose" [Gaurce, 2010]. The registration view of the VGI_Editor application is a non-mapping page that provides users with a registration form to fill in to have a full access to the VGI_Editor features. A user needs to provide the following information: username, e-mail address and password, as shown in Figure 3.10.

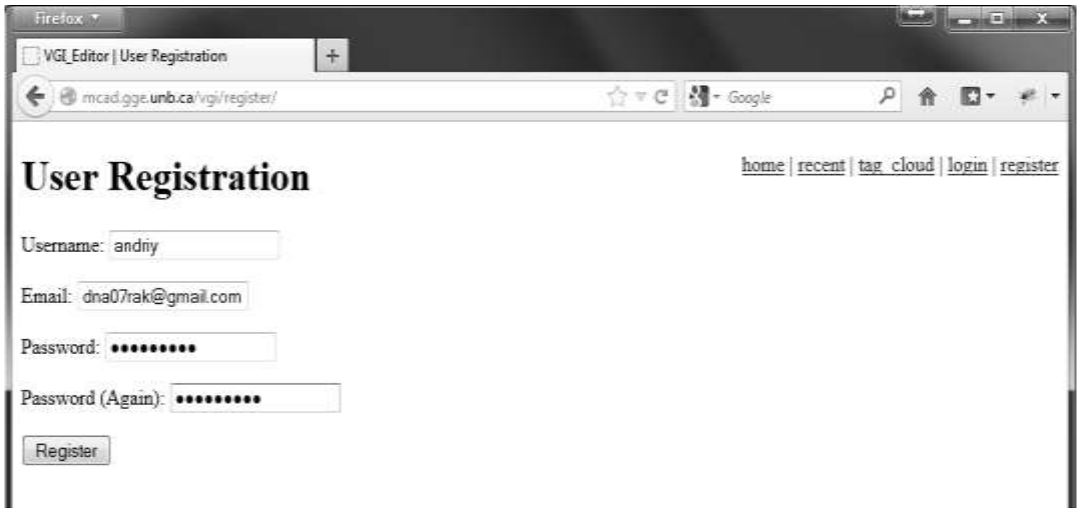


Figure 3.10 User Registration Form

Research led by Nkhwanana [2009] suggests that a registration system should be established in place when validating VGI. Such registration system should not compromise contributors' privacy and/or security, but should rather allow "users to be identified only for the purposes of reliably identifying trustworthy contributions" [Nkhwanana, 2009]. Information required from the users on the VGI_Editor website provides enough details to be able to identify a user, but will not violate contributors' privacy and/or security. Similar registration systems are in place for two most prominent VGI projects: OpenStreetMap and Wikimapia, and is shown in Figures 3.11 and 3.12, respectfully.

Figure 3.11 OpenStreetMap Project Registration Page

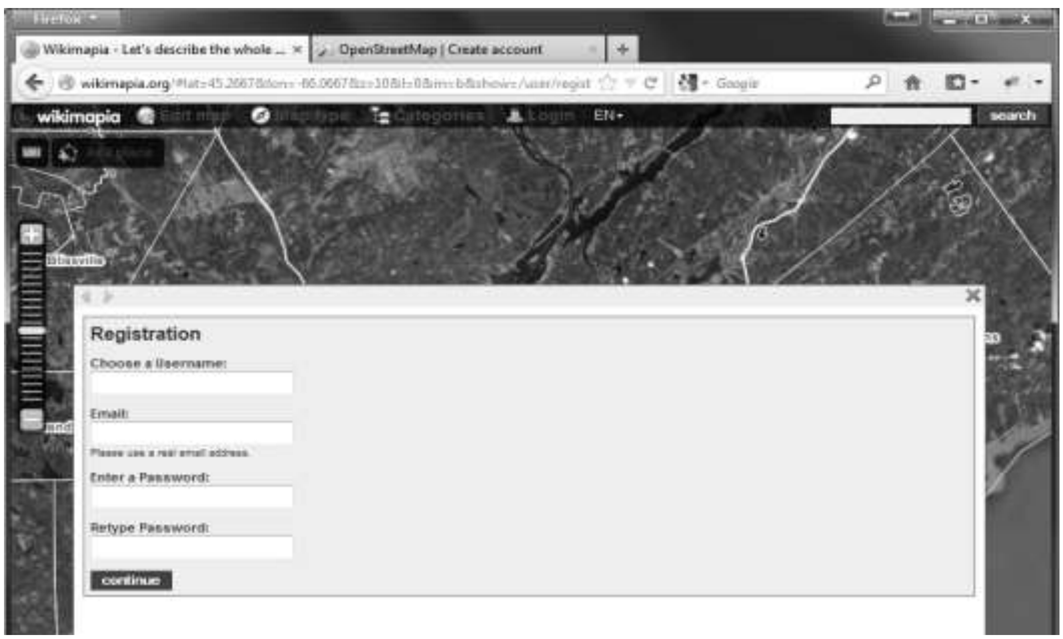


Figure 3.12 Wikimapia Project Registration Page

After successful registration within the VGI_Editor application, the user will be able to login and logout from the system, as shown in Figure 3.13.

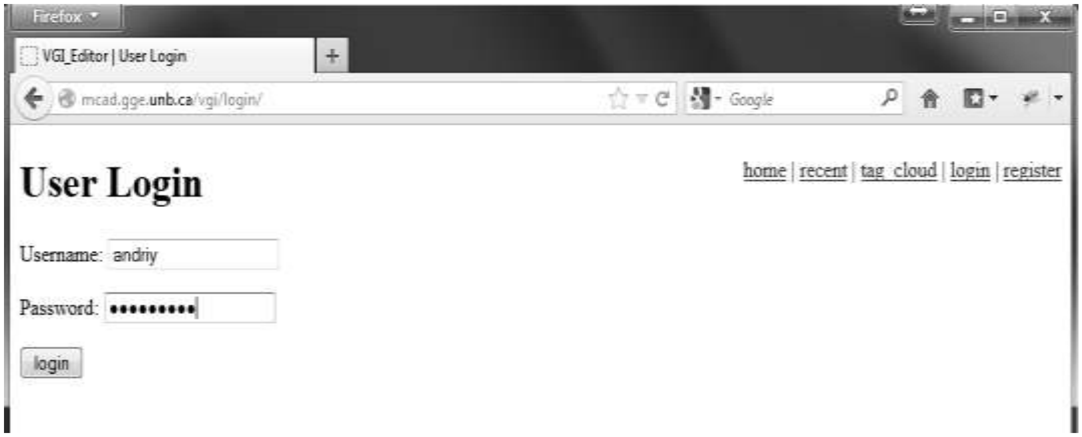


Figure 3.13 VGI_Editor Login View

3.4.3 VGI Contribution View

The contribution page is a non-mapping page that provides users with tools necessary to upload VGI in the form of a shapefile (to be discussed in more detail in section 3.5), and is shown in the Figure 3.14. Registered users willing to submit VGI must follow these following steps:

- Have a shapefile archived in a *.zip* format, comprised of the following mandatory files:
  - "filename.shp – main file that provides access to a variable-record-length file in which each record describes a shape with a list of its vertices

- filename.shx – index file, in which each record contains an offset of the corresponding main file record from the beginning of the main file
- filename.dbf – a table file that contains feature attributes with one record per feature, providing a one-to-one relationship between geometry and attributes
- filename.prj – a projection definition file, stores coordinate system information" [Digitalpreservation, 2011].

- Provide a character encoding of the file.
- Provide a title/name for VGI.
- Use tags as a means to describe VGI (more details in subsection 3.4.4). And
- Decide whether to share VGI with the VGI_Editor community or keep it restricted only to users and their list of friends (more details in subsection 3.4.5).

Figure 3.14 VGI_Editor Contribution Page

3.4.4 Tag Cloud View

A tag is a "keyword used to describe an article or website", and tag cloud is a "box containing a list of tags with the most prominent or popular tags receiving a darker and bigger font than less popular tags" [About.com, 2012]. For the purposes of this research a tag will represent a keyword used to describe different aspects of VGI, such as accuracy, methods used for collecting, processing, storing VGI, scale, etc.

A "Tag cloud" view serves as the means of expressing metadata on contributed VGI. "Metadata plays a key role in facilitating access to up-to-date spatial information and contributes to the finding and delivering of high quality spatial information services to users" [Kalantari et al., 2010]. The concept of utilizing tagging and tag cloud as a means of metadata has been described by Kalantari et al. [2010] and states:

> *The tagging approach provides more freedom for users, because when tagging, the user does not have to make a decision and restrict the resource to just one or two formal terms from a controlled keywords they may or may not be familiar with. Instead the users supply their own terms which are meaningful to them. For instance, spatial data users can select different tags to describe the same item. Items related to scale may be tagged '500', '1:500', '1/500'. This flexibility allows users to classify their collections of items in a way that they find useful and users also have to decide whether each tagged item is actually relevant to what they're looking for.*

Berendt and Hanser [2007] argue that "tags are not metadata, but 'just more content'". While this statement could be true for GI coming from authoritative datasets created by highly-trained professionals, this research assumes that tagging and use of a tag cloud could replace creation of metadata for VGI. As in the situation where most VGI comes with very little to no metadata information, a tagging system could provide very valuable information about contributed VGI. Although this approach is not compliant to OGC standards and doesn't fit with CGDI approach, it is seen as a useful tool in situations where metadata information is not accompanying VGI. A tag cloud page of VGI_Editor is shown in Figure 3.15.

Figure 3.15 VGI_Editor Tag Cloud Page

3.4.5 User View

The user view is a non-mapping page that provides users with access to the VGI submitted by other users and also to VGI submitted by their friends, and is shown in Figure 3.16.

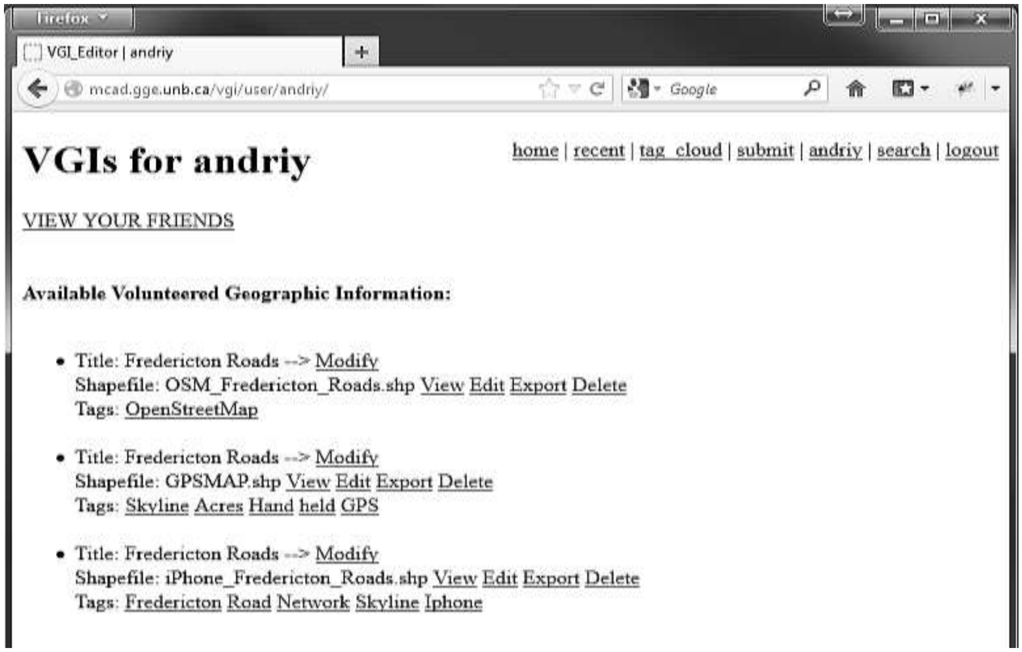


Figure 3.16 VGI_Editor User View

From this view, a user can exercise the following functions of the prototype:

- View, edit, export and delete VGI from the VGI_Editor application. The function to delete VGI from the VGI_Editor has only the user who contributed VGI, thus ensuring that users maintain and can exercise their Intellectual Property (IP) rights to VGI. Such IP rights of a user are not violated by other users, since, as mentioned in subsection 3.4.3, users have the ability to choose between making their contributions public and editable by all registered users of the program, or keeping it private, thus making it visible only to themselves and their friends.
- Edit information on contributed VGI, as shown in Figure 3.17. Users can change the title of the contributed VGI together with tags associated with

VGI. User can also 'share' such VGI with all registered VGI_Editor users, if he/she has not done so previously.

Figure 3.17 Editing Information on VGI

- Create a "circle of friends" and invite new friends to join the VGI_Editor application, as shown in Figure 3.18. Through the VGI_Editor, users can invite new friends over e-mail to join the VGI_Editor application. Users need to provide an e-mail address of the friend, and the application will automatically generate and send an invitation, as shown in Figure 3.19. Users can also create their own "circle of friends" using the VGI_Editor application by adding already registered users. In this case such a circle can work together on contributed VGI, excluding all other users not belonging to such a group from making modifications. This tool can be useful if users are worried about the possibility of harmful content being introduced into their VGI by users they do not know, or for example by users outside their own company, and as such working together with a group of ‘friends’ or co-workers they trust and with whom they are willing to work.

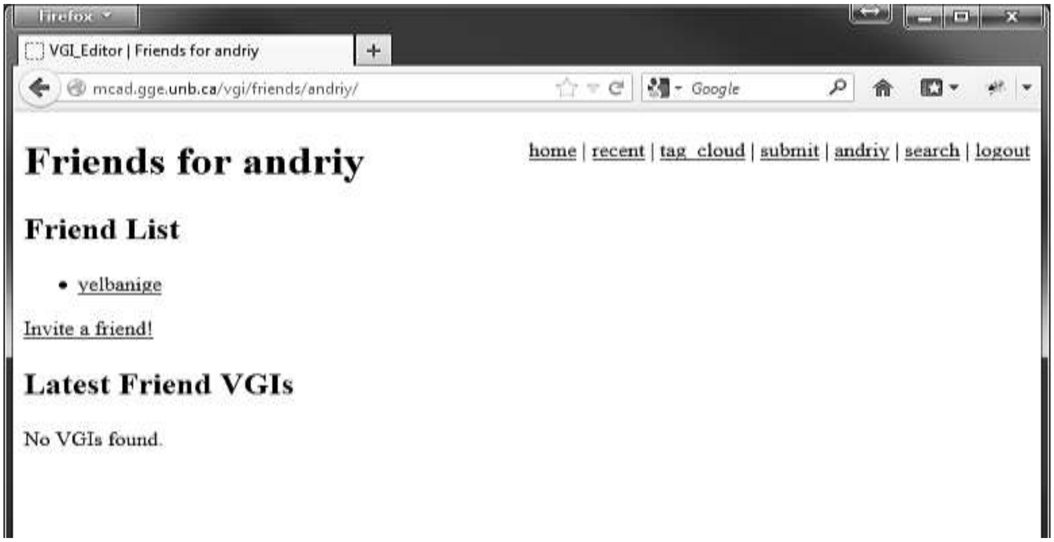


Figure 3.18 VGI_Editor Friends Circle

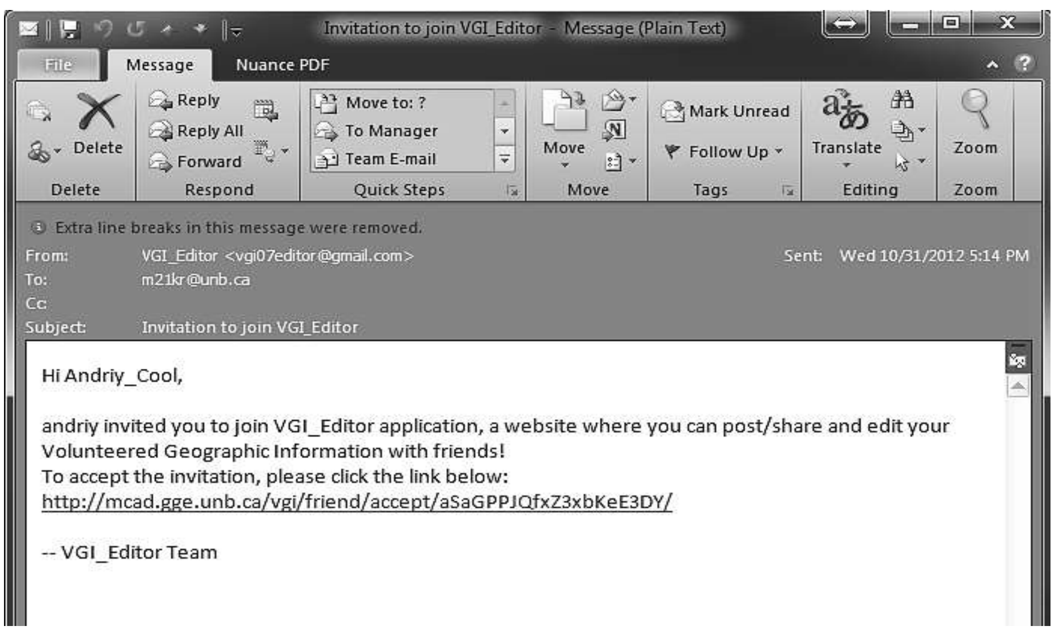


Figure 3.19 VGI_Editor Invitation E-Mail

3.4.6 VGI View

The VGI view consists of two mapping (view and edit) and one non-mapping (comments) pages. The VGI view supplies users with tools essential to validate contributed VGI to comply with CGDI accuracy standards. Validation process encompasses both crowdsourcing and authoritative methods of VGI validation.

The first mapping page, as shown in Figure 3.20, provides users, both registered and non-registered, with the ability to view contributed VGI without making any changes. Such VGI could be overlaid on top of a number of baselayers to better visualize and understand VGI. Baselayers available to users of VGI_Editor include:

- Google Roadmap layer – displays the road map view;
- Google Satellite layer – displays Google Earth satellite view;
- Google Hybrid layer – displays a mixture of normal and satellite views;
- Google Terrain layer – displays a physical map based on terrain information [Google, 2012];
- OpenStreetMap baselayer;
- National Road Network (NRN) from GeoBase portal – only part for city of Fredericton, New Brunswick; and
- Vector layers (VGI) contributed directly by the users, their circle of friends and VGIs made public by other users.

Figure 3.20 VGI Overlaid on Top of Google Roadmap Layer

The second mapping page provides registered users with the ability to view and edit both geometry and attributes of contributed VGI. The third VGI page - comments page together with edit mapping page - are both examples of crowdsourcing validation of contributed VGI and will be discussed in subsection 3.4.6.1.

3.4.6.1 Crowdsourcing Validation of VGI

In this research, a crowdsourcing validation of VGI describes activities directed towards validation and/or modification of VGI to comply with CGDI accuracy standards.

The VGI_Editor, based on techniques described in Section 2.2, employs the following crowdsourcing validation techniques in its design:

- Modification of geometry and attributes of VGI by registered users,
- Voting system, and
- Comments.

The above mentioned crowdsourcing techniques have been discussed and proposed by Nkhwanana [2009] in his research as means of VGI validation based on their use and popularity in the following projects:

- Wikipedia (work of multiple people to validate and/or change data);
- Facebook (simple like/dislike voting system, to choose the content that is the most important to the users);
- Amazon and eBay (comments are valuable source of information to users who consider making an on-line transaction, and give more confidence to the buyer) [Nkhwanana, 2009].

The edit mapping page provides access to the first crowdsourcing validation technique – modification of geometry, Figure 3.21, and/or attributes, Figure 3.22, of contributed VGI.

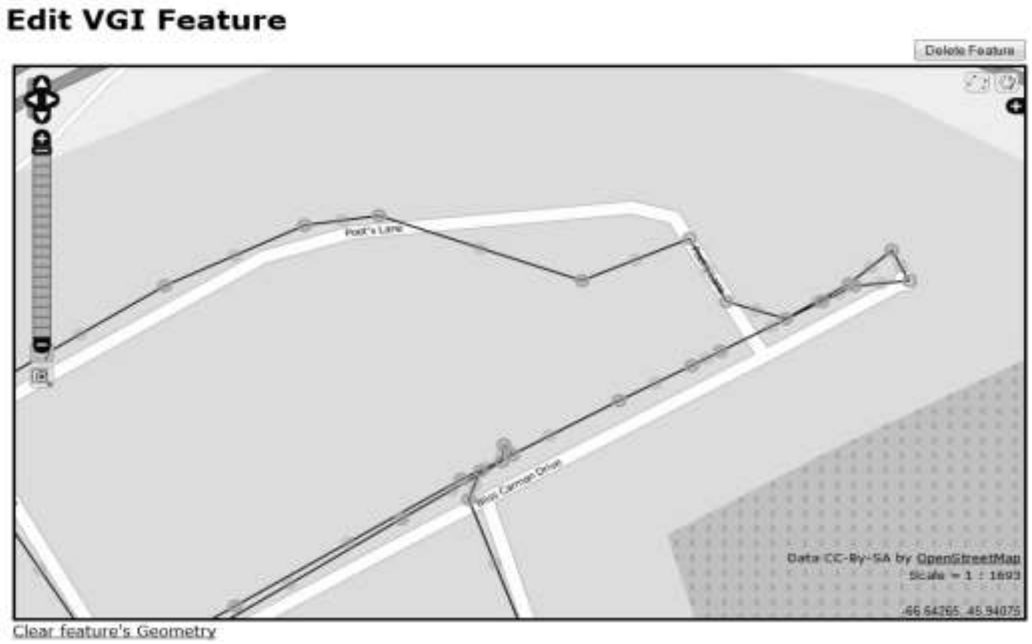


Figure 3.21 Modification of VGI Geometry Based on OpenStreetMap Baselayer

Figure 3.22 Modification of VGI Attributes

The modification of VGI geometry is a core tool in VGI_Editor application. Using this tool, baselayers and users local knowledge of the area, volunteers can modify the geometry of VGI to represent their best "estimate" of the features positional accuracy. To improve the positional accuracy of VGI, users should have an understanding of baselayers data quality. While there is no data quality information on baselayers provided by Google and OpenStreetMap, data quality information on National Road Network (GeoBase NRN v.2 validated in 2003) from GeoBase portal is as following [GeoBase, 2012b]:

- Attribute accuracy - the method used for evaluating the accuracy of the non-quantitative attribute values with respect to reality is determined by each data provider.
- Completeness - NRN road representation corresponds to centerline of all non-restricted usage roads (5 meters wide or more, drivable and not blocked by an obstacle). Roads isolated from the main road network, alleyways, resource and recreational roads and addressing information may not be included in some NRN datasets.
- Horizontal positional accuracy - the planimetric accuracy aimed for the product is 10 m or better. Under the data maintenance phase, no systematic validation of geometric and attributive accuracies is performed on all NRN datasets.

One more technique to improve the quality of VGI is to have multiple baselayers with which to compare VGI. In cases where VGI contains GI that is not represented in a baselayer, use of VGI collected for the same area by other volunteers could be used. For example, in Figure 3.23, where a solid black line represents the NRN baselayer and solid grey line represents VGI collected using GPSDrive™ application software on an iPhone, there is an area (highlighted with black box) where there is no corresponding GI stored in the NRN baselayer.

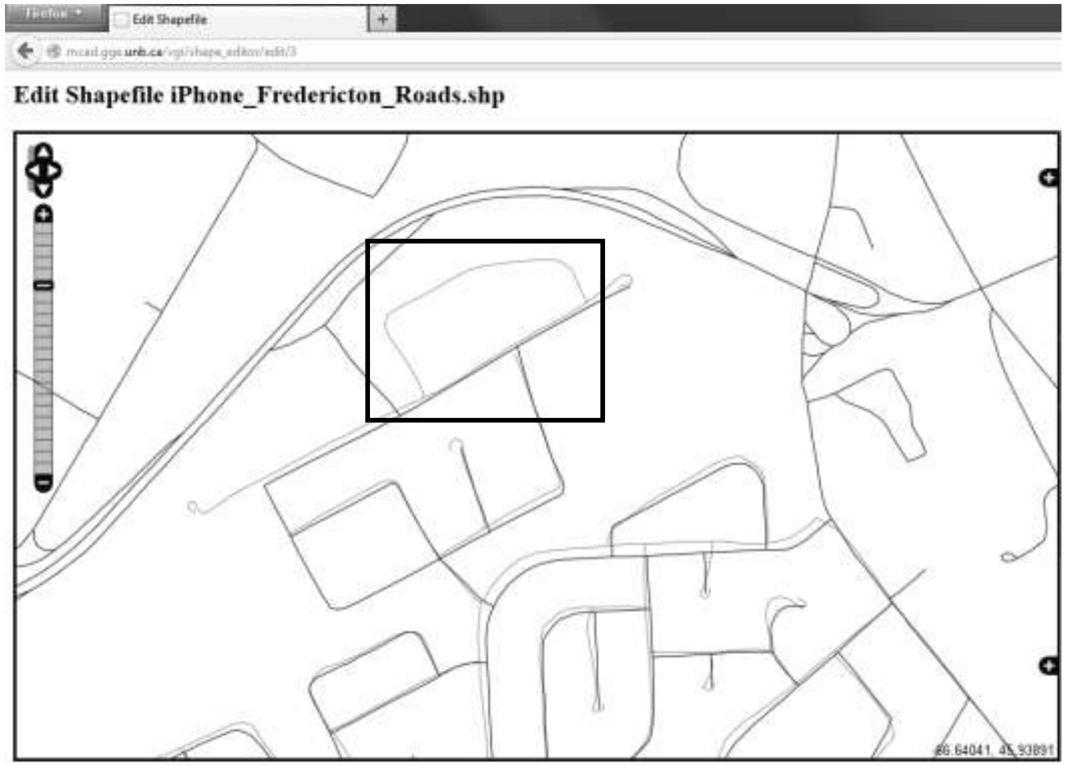


Figure 3.23 VGI Collected Using iPhone Overlaid with NRN Baselayer

In such a case VGIs collected by other users could be used as a baselayer to validate its quality. In Figure 3.24, where a solid black line represents VGI collected using GPSMAP handheld GPS and solid grey line represents VGI collected using iPhone. Note that both VGI layers include a road that is not included within NRN baselayer. As both layers represent the same road almost identically, VGI collected using iPhone could potentially be used to update the missing part of the road in NRN baselayer.

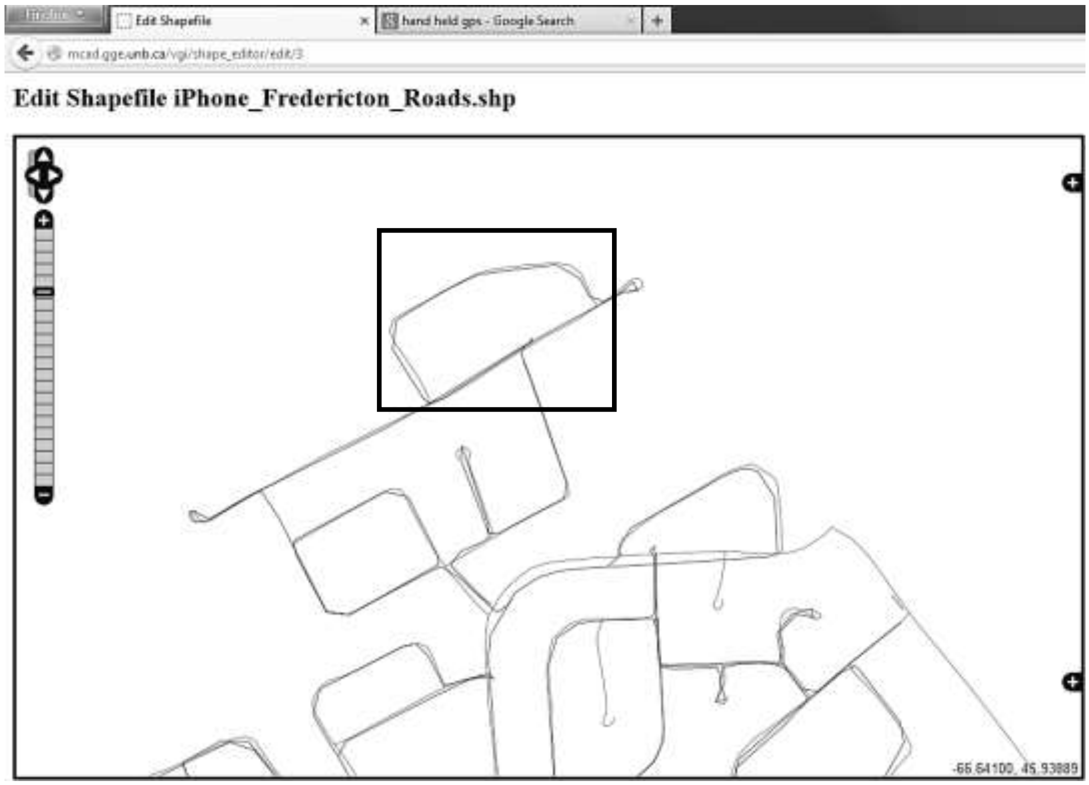


Figure 3.24 VGI collected Using iPhone Overlaid with VGI Collected Using GPSMAP Handheld GPS

In addition to providing tools for modifying geometry of VGI, the VGI_Editor edit mapping page also provides tools for validating attribute accuracy of VGI. Such validation is performed through crowdsourcing work of VGI_Editor registered users on modifying existing attributes of VGI (Figure 3.25) and populating such VGI with new attributes (Figure 3.22).

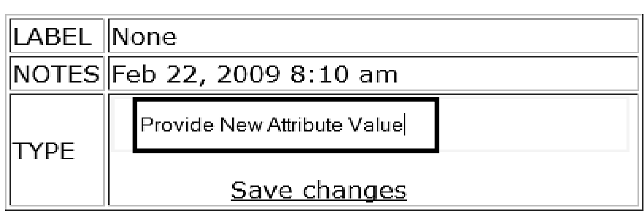


Figure 3.25 VGI_Editor Modifications of Attribute Value

The third VGI page - Comments - provides access to voting and comments crowdsourcing validation techniques (Figure 3.26). On this page, registered users can employ a simple "Like/ Dislike" voting system by clicking on the "thumb up" or "thumb down" image to express their opinion whether or not such VGI is considered suitable for their needs. The similar voting system is currently utilized in Facebook, YouTube, and most eCommerce websites. Based on the number of likes and dislikes users have a general idea whether to use such information. The comments section of the VGI page offers users a chance to leave feedback on VGI, its accuracy, users who contributed to it, etc.

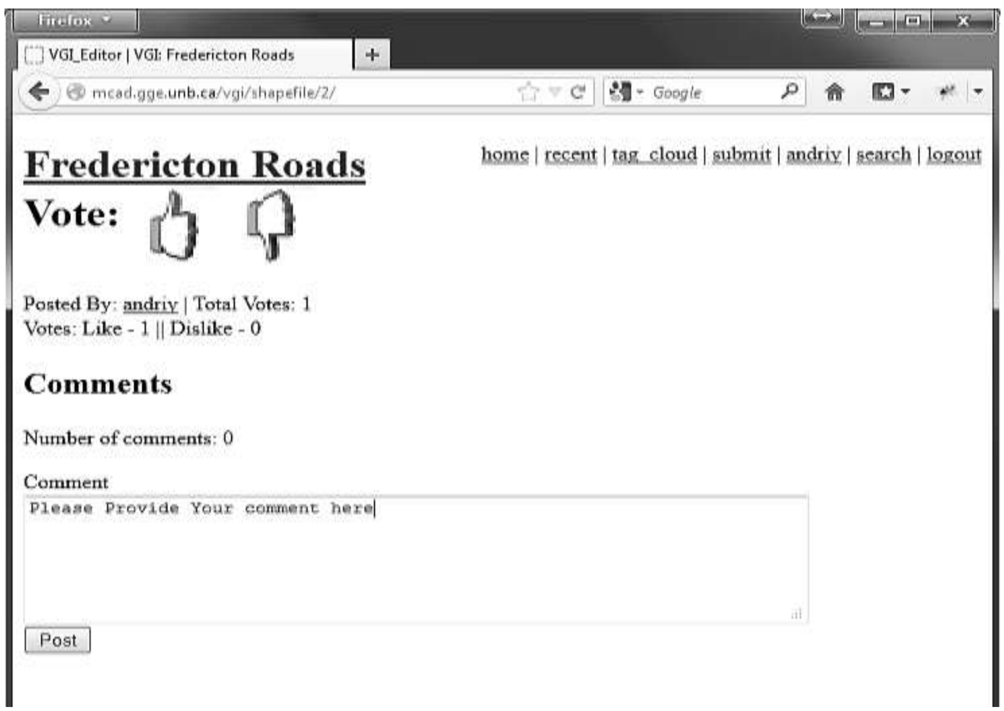


Figure 3.26 VGI_Editor Comments Page

3.4.6.2 Authoritative Validation of VGI

Authoritative validation of VGI is based on the Goodchild and Hunter [1997] methodology for evaluating positional accuracy – Buffer Overlay Method (BOM). This method is only one of the methods available to validate positional accuracy of VGI, and can't be used validate other components of VGI data quality. As such more research is needed to define a better method for this purpose. "The buffer comparison method was selected because it was easier and faster to use for analyzing large or multiple datasets" [Sabone, 2009]. Positional accuracy of VGI is estimated by measuring the difference between VGI and the NRN dataset - a data source of higher accuracy.

3.4.7 VGI_Editor Legal Liability Considerations and Functions

A core function of the VGI_Editor prototype is the incorporation of RMTs, described in Section 2.7, necessary to handle legal liability issues surrounding implementation of VGI with CGDI. In this section, four primary risk management techniques are recommended to CGDI in order to manage their legal liability risks. The first RMT is identification of possible risks, which includes the following steps:

1. Establishment of a risk management team;
2. Identification and ranking by the team of appreciable risks originating from integration of VGI into CGDI datasets;

3. Evaluation by the team of the likelihood of risk occurrence against the magnitude of harm;
4. Development by the team of processes/stages to minimize/reduce the most important risks;
5. Application of selected processes/stages to identified risks;
6. Conducting regular checks on samples of data, to verify if risks have been solved or if they are still present;
7. Continuous identification of new risks by the risk management team, which involves following procedures starting from step 2.

Within this RMT context, the VGI_Editor can be considered to be an example of a "Step 4" item. It is a coherent protocol of processes and stages applied to the CGDI risks to ensure that accuracy of VGI is satisfying CGDI data quality requirements and thus limit potential liability for CGDI.

A second proposed RMT suggests keeping constant focus on quality assurance of GI, including incorporated VGI. This RMT is based on the argument that high quality, accurate GI is less prone to cause legal liability issues than data of less accuracy. The VGI_Editor provides tools to perform quality checks on VGI described in subsections 3.4.6.1 and 3.4.6.2.

The third RMT technique proposed in this research includes the use of disclaimers in contracts through End User License Agreements (EULAs). At GeoBase [2012b] portal it is stated that "all distributed GeoBase NRN datasets should be accessed and used relatively to the GeoBase Unrestricted Use Licence Agreement", which indicates the following concerning legal liability:

> *Canada makes no representation or warranty of any kind with respect to the accuracy, usefulness, novelty, validity, scope, completeness or currency of the Data and expressly disclaims any implied warranty of merchantability or fitness for a particular purpose of the Data. Canada does not ensure or warrant compatibility with past, current or future versions of your browser to access the site's Data.*
>
> *The Licensee shall have no recourse against Canada, whether by way of any suit or action, for any loss, liability, damage or cost that the Licensee may suffer or incur at any time, by reason of the Licensee's possession or use of the Data.*
>
> [GeoBase, 2012a]

The last RMT proposed in this research suggests the inclusion of a duty to warn about the quality of VGI. At a minimum the GI provider should always exercise a duty to warn the user/receiver of the information about the quality of the data and indicate any limitations that might come with it [Chandler and Levitt, 2011]. Additional information on a duty to warn responsibility of GI providers can be found in the works of Gervais [2003], Bédard et al. [2004], and Levesque et al. [2007]. The liability disclaimer should clearly point out that the volunteered information is to be used only for certain purposes, or be employed at the user's own risk [Club TomTom, 2007]. The VGI_Editor is designed to include such disclaimer both when data is downloaded from the website, as shown in Figure 3.27, and also after such download occurred. The liability disclaimer in Figure 3.27 example doesn't correspond to a warning as defined by ISO standards, and would require some modifications. Such information on data quality can be passed along to users using their e-mail addresses, which are recorded during the registration process to the VGI_Editor. The above mentioned risk management techniques are directed towards preventing CGDI from exposure to legal liability arising from integration of VGI.

!!! Please consult an expert before using selected VGI !!!

The VGI_Editor team makes no representations or warranties as to the accuracy of the information presented and the user assumes the entire risk as to the use of any or all information.

It is hard for VGI_Editor team to foresee potential users and their requirements to Quality of the Data. It is recommended for You to make an informed decision on the data quality based on the tools provided in this application, and make a judgement regarding fitness of such data for Your intended use. In case it is hard to make an informed judgement please consult an expert before using selected VGI.

Figure 3.27 VGI_Editor Disclaimer on Data Quality

3.5 Spatial Data Format

VGI_Editor employs a proprietary *.shp* or "Shapefile" spatial data format to represent VGI contributions. Shapefiles are commonly used for “sharing vector data between GIS applications” [Rivix, 2012] and, as such, almost any spatial data format can be converted into a Shapefile. When a user imports a Shapefile, its content is stored into the database using the Well Known Text format so that other users can work with it. A generic representation of Shapefile inside the database is shown in Figure 3.28, and a more detailed representation is shown in Figure 3.29.

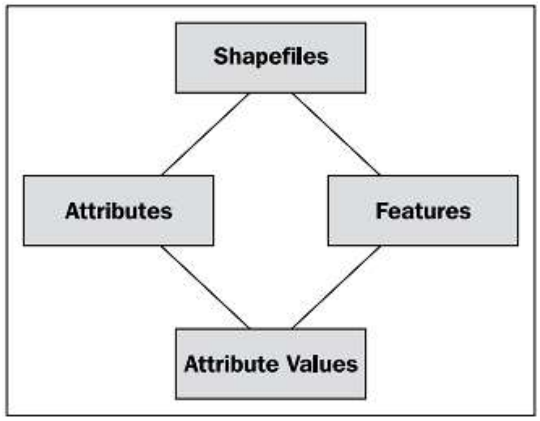


Figure 3.28 Generic Representation of Shapefile format Inside the VGI_Editor Database (from Westra [2010])

Each imported VGI will be represented by:

> *a single Shapefile object in the database. Each Shapefile object will have a set of Attribute objects, which define the name and data type for each attribute within the Shapefile. The Shapefile object will also have a set of Feature objects, one for each imported feature. The Feature object will hold the feature's geometry, and will in turn have a set of AttributeValue objects holding the value of each attribute for that feature.*
>
> [Westra, 2010]

To be able to work simultaneously with all contributed VGIs, such information should be converted into the same spatial reference system. To accomplish this, all Shapefiles imported into VGI_Editor are converted into latitude and longitude coordinates referenced to the World Geodetic System (WGS) 84 datum. When VGI is overlaid on top of base layers provided by VGI_Editor, such information is reprojected on the “fly” to match the spatial reference of the baselayer.

```
class Shapefile(models.Model) :
    filename = models.CharField(max_length = 400)
    srs_wkt = models.CharField(max_length = 700)
    geom_type = models.CharField(max_length = 50)
    encoding = models.CharField(max_length = 20)
    title = models.CharField(max_length = 200)
    user = models.ForeignKey(User, editable = False)
    submitted_date = models.DateTimeField(default=datetime.datetime.now)
    objects = models.GeoManager()

    def __unicode__(self) :
        return self.filename

class Attribute(models.Model) :
    shapefile = models.ForeignKey(Shapefile)
    name = models.CharField(max_length = 255)
    type = models.IntegerField()
    width = models.IntegerField()
    precision = models.IntegerField()
    objects = models.GeoManager()

    def __unicode__(self) :
        return self.name

class Feature(models.Model) : # Features inside database are in SRID==4326
    shapefile = models.ForeignKey(Shapefile)
    geom_point = models.PointField(srid = int(settings.FEATURE_SRID),
                                   spatial_index = True, blank = True, null = True)
    geom_multipoint = models.MultiPointField(srid = int(settings.FEATURE_SRID),
                                             spatial_index = True, blank = True, null = True)
    geom_multilinestring = models.MultiLineStringField(srid = int(settings.FEATURE_SRID),
                                                       spatial_index = True, blank = True, null = True)
    geom_multipolygon = models.MultiPolygonField(srid = int(settings.FEATURE_SRID),
                                                 spatial_index = True, blank = True, null = True)
    geom_geometrycollection = models.GeometryCollectionField(srid = int(settings.FEATURE_SRID),
                                                             spatial_index = True, blank = True, null = True)
    objects = models.GeoManager()

    def __unicode__(self) :
        return str(self.id)

class AttributeValue(models.Model) :
    feature = models.ForeignKey(Feature)
    attribute = models.ForeignKey(Attribute)
    value = models.CharField(max_length = 255,
                             blank = True, null = True)

    def __unicode__(self) :
        return unicode(self.value)
```

Figure 3.29 Detailed Representation of Shapefile format (Model) Inside the VGI_Editor Database

3.6 Summary

This chapter described the implementation of the VGI_Editor prototype design. The design included discussions on VGI_Editor architecture, choice of web application framework, and client side implementation. Key concepts were discussed, and the chapter described the implementation of crowdsourcing and authoritative techniques for validating VGI, and implementation of RMTs to limit/reduce potential legal liability issues for CGDI.

The chapter concluded with descriptions of the spatial data format used by the VGI_Editor and, where applicable, the reprojection of VGI contributions to match the spatial reference of baselayers provided by VGI_Editor prototype. The next chapter evaluates the results of testing VGI using the VGI_Editor.

4. Testing and Evaluation of the VGI_Editor Prototype

The preceding chapters described the design and implementation of the VGI_Editor prototype aimed at validation of VGI to comply with CGDI accuracy standards. This chapter discusses the suitability of using VGI in CGDI datasets by examining performance of VGI against CGDI datasets. The evaluation consisted of testing VGI against CGDI dataset based on the following measures: positional accuracy, attribute accuracy, data structure, metadata and uncertainty.

4.1 Prototype Evaluation

In Section 1.4 set of evaluation criteria for the prototype has been established and includes the following:

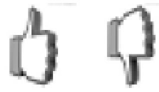
1. Satisfaction of the CGDI accuracy standards by VGI after being validated through the VGI_Editor prototype; and
2. Incorporation of tools within the prototype to limit/reduce potential legal liability from using VGI.

Essentially, the criteria above denote that the VGI_Editor prototype will be deemed "successful" if the tools it provides are capable of validating VGI to comply with CGDI accuracy standards; and to limit/reduce potential legal liability from such integration of datasets. The prototype itself was not tested; instead VGI that has been validated using the VGI_Editor prototype was tested. Such VGI has been tested to see whether it meets CGDI accuracy requirements or not, and is described below in Sections 4.2 and 4.3.

VGI was validated by users employing the VGI_Editor prototype. The users were asked to validate VGI datasets (description of the datasets is provided in Section 4.2) uploaded into the VGI_Editor prototype. The users were requested to validate the VGI using the tools within the VGI_Editor prototype that have been discussed in Chapter 3. Users were given a half an hour presentation explaining the concepts of VGI, different approaches to its validation, legal issues that might arise from using VGI, and the benefits of its integration with authoritative datasets. The presentation was followed by step-by-step guidance on how to use the VGI_Editor prototype to validate VGI using crowdsourcing validation techniques described in Subsection 3.4.6.1. After the demonstration the users were asked to validate VGI on their own free time by using the VGI_Editor prototype. Users were represented by: one undergraduate student taking the GGE 2423 *Introduction to Geographic Information Systems* course and nine graduate students in GIS/Remote Sensing Stream of the Department of Geodesy and Geomatics Engineering at the University of New Brunswick. Users understood what was expected of them, and did not have any major problems with the software. Figures 4.1 and 4.2 present comments provided by the users on their experience with the VGI_Editor prototype.

Fredericton Roads

Vote:



Posted By: andriy | Total Votes: 6
Votes: Like - 4 || Dislike - 2

Comments

u740u said:

great!

Nov. 15, 2012, 12:42 p.m.

131.202.11.33

AndyPat said:

Like the app but made a mistake!

Nov. 15, 2012, 1:16 p.m.

131.202.9.36

d29pq said:

Great job, nice mapping tool

Nov. 19, 2012, 5:38 p.m.

131.202.161.64

yelbanige said:

Nice prototype, needs additional functionality in production. The rough edges can be improved. Good Job.

Nov. 26, 2012, 11:51 a.m.

131.202.8.137

Number of comments: 4

Figure 4.1 Comments Provided by Users of the VGI_Editor Prototype on VGI collected using GPSMAP handheld GPS

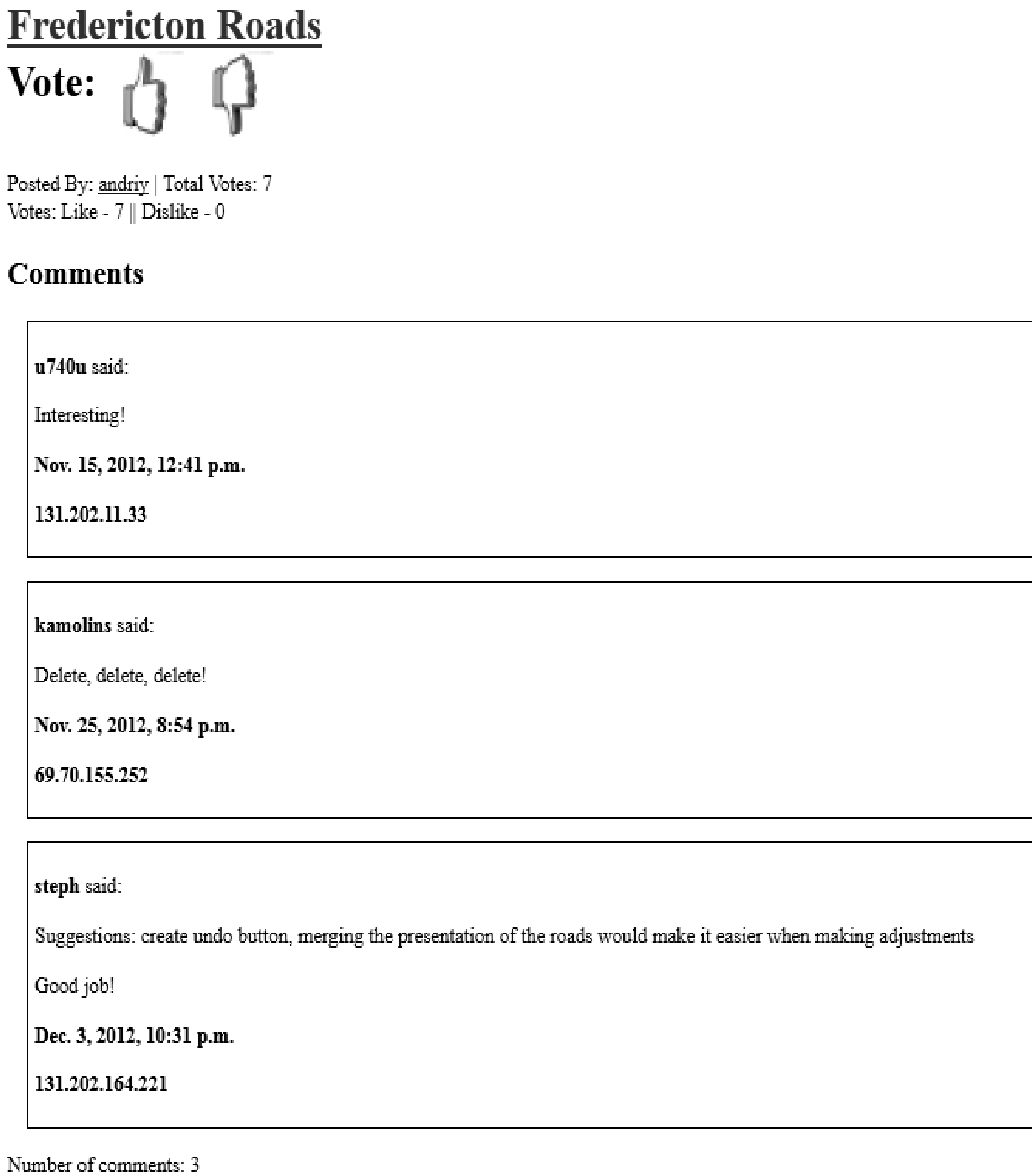


Figure 4.2 Comments Provided by Users of the VGI_Editor Prototype on VGI collected using iPhone

The results of VGI validation by the users of the VGI_Editor prototype are presented in the Subsections 4.2.1 and 4.2.2.

4.2 VGI Validation

The two VGI datasets came from field measurements collected by Sabone [2009], and represent residential streets in Skyline Acres, a subdivision within the city of Fredericton, New Brunswick, Canada (Figure 4.1). The streets were captured as individual separate GPS tracks using an Apple iPhone equipped with GPS MotionX software and a Garmin 76CSx handheld GPS. Later, these separate GPS tracks were merged together to create two VGI datasets: iPhone VGI and GPSMAP VGI.

Figure 4.3 VGI Collected Area

The VGI datasets were compared to an authoritative dataset – NRN [GeoBase, 2012b] – for the City of Fredericton, from GeoBase portal, using an authoritative validation method described in Subsection 3.4.6.2.

In the original research design, an authoritative validation method - BOM - was intended to be a part of the VGI_Editor prototype. Due to time constraints the BOM tool was not implemented in the VGI_Editor prototype; instead, ESRI ArcGIS software was used to evaluate the VGI contributions to check whether such datasets satisfied the requirements for being included as a source of updates for CGDI datasets. Two main components for evaluation of VGI were established using ESRI's ArcGIS software:

- A personal Geodatabase Feature Dataset with defined coordinate system – NAD_1983_CSRS_New_Brunswick_Stereographic (Figure 4.4) – were created to ensure that both VGI and NRN datasets have the same coordinate system, and to assist with evaluation by making it possible to perform measurements using metric units rather than geographic coordinates.

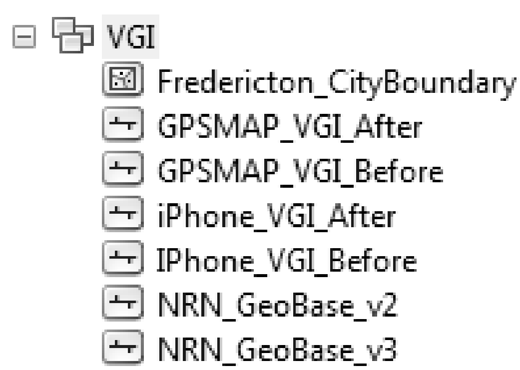


Figure 4.4 A Personal Geodatabase Feature Dataset for Storing VGI and NRN Data

- A BOM Toolbox (Figure 4.5.) that provides tools for evaluating positional accuracy and uncertainty of VGI.

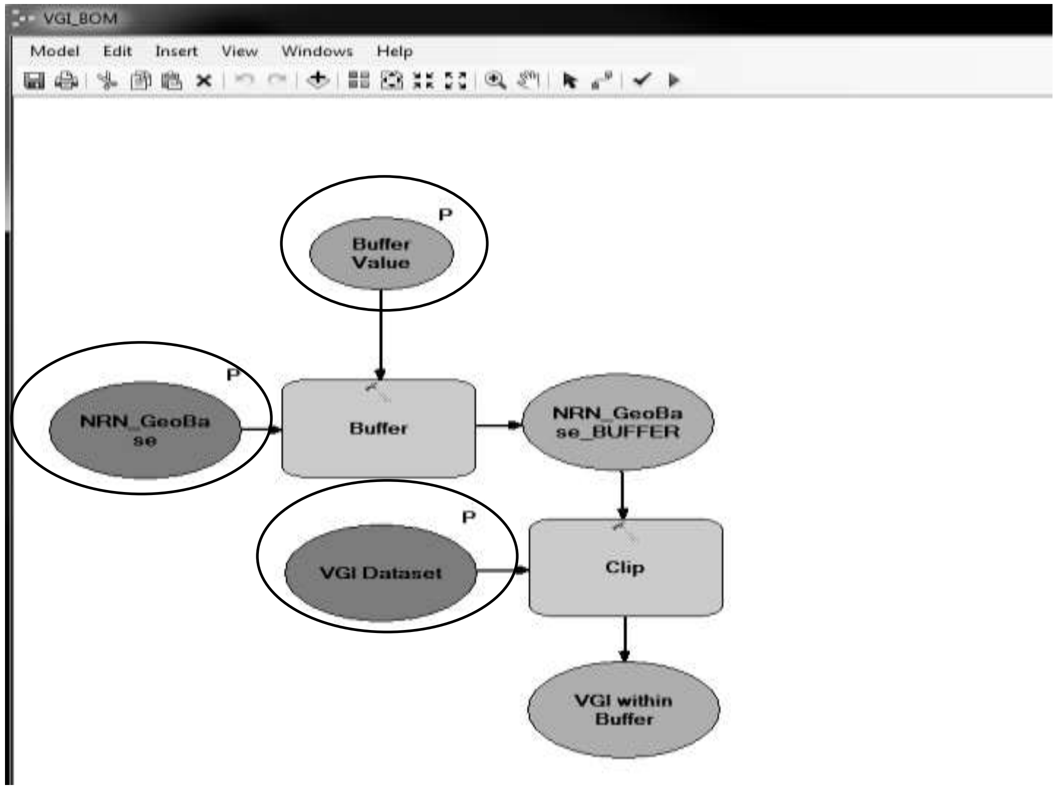


Figure 4.5 BOM Toolbox, built in ESRI ArcGIS Software, for Validating Positional Accuracy and Uncertainty of VGI

The Toolbox takes three parameters, identified with letter P in the Figure 4.5. The first parameter comprises the authoritative dataset, the second one represents the distance around the first parameter (NRN dataset) in metres to create the buffer zone. The last parameter represents the VGI dataset that will be evaluated. The comparison was carried out using the following steps [Sabone, 2009], as shown in Figure 3.27:

- Create a layer with a 10 m buffer (second parameter) around NRN streets (first parameter) (the buffer distance of 10 m is derived from the positional accuracy criterion for integrating VGI into CGDI datasets, which stipulates: VGI must have a planimetric accuracy of ± 10 m for roads);

- Intersect the buffer layer with VGI (third parameter) streets data;
- *The total length of intersected road track data in this layer was then divided by the total length of the original VGI street before clipping (the VGI and NRN datasets were ensured to have the same extent), and then multiplied by a hundred to give a percentage value. The equation used is shown below:*
  *(Result of Intersection Layer total length ÷ Original Layer total length) * 100*

  [Sabone, 2009].

To evaluate one aspect of the VGI_Editor prototype's functionality that provides tools to validate VGI to comply with CGDI accuracy standards, the VGI datasets were compared against NRN data based on criteria established for integration VGI and CGDI datasets twice:

1. First, the VGI datasets were evaluated against NRN data in their original/unmodified form; and
2. Secondly, the VGI were compared against NRN data after being validated by the users of the VGI_Editor prototype, described in Section 4.1.

Comparing results of the VGI validation both in its raw form and after being assessed by users of the prototype indicates whether the research objective "design and deploy a prototype that is capable of validating VGI to comply with CGDI positional accuracy standards in particular" has been met.

The following datasets were used in this research:

1. iPhone_VGI_Before – VGI dataset collected by Sabone [2009] from the field on February 22, 2009 using an iPhone device;

2. iPhone_VGI_After – VGI dataset collected by Sabone [2009] from the field on February 22, 2009 using an iPhone device, which was later validated through the VGI_Editor prototype, using tools described in Subsection 3.4.6 and Section 4.1;
3. GPSMAP_VGI_Before – VGI dataset collected by Sabone [2009] from the field on February 22, 2009 using a Garmin GPSMAP 76CSx handheld GPS;
4. GPSMAP_VGI_After – VGI dataset collected by Sabone [2009] from the field on February 22, 2009 using a Garmin GPSMAP 76CSx handheld GPS, which was later validated through the VGI_Editor prototype, using tools described in Subsection 3.4.6 and Section 4.1;
5. NRN_GeoBase_v3 – GeoBase NRN edition 3 validated in 2011, that was used as an instance of CGDI dataset.

VGI was validated using the criteria established in Section 2.3 for positional, attribute accuracy and their uncertainty, data structure and metadata.

4.2.1 Positional, Attribute Accuracy and Their Uncertainty

To satisfy the objective of research to design and deploy a prototype that is capable of validating VGI to comply with CGDI accuracy standards, to be "successful" the criteria mentioned in Section 2.3 should be met by VGI validated using the prototype. The BOM is used to check the requirements for positional accuracy and uncertainty associated with it. The requirements for positional accuracy stipulates that VGI must have a planimetric accuracy of no less than ±10 m for roads (in clear unobscured areas), and ± 30 m or less for other features (in clear unobscured areas). The uncertainty associated with such accuracy specifies that 90% or more of all well-defined features in a dataset should be within a specified distance (in this case 10 m) of their true position. Table 4.1 shows the results of comparing VGI datasets with the GeoBase NRN data.

Table 4.1 Results of Testing VGI Positional Accuracy

VGI Dataset	**Percentage Within 10m of NRN dataset**	**Change of Accuracy**
iPhone_VGI_Before	82.89%	9.23%
iPhone_VGI_After	92.12%	
GPSMAP_VGI_Before	91.11%	4.24%
GPSMAP_VGI_After	95.35%	

Results of comparing positional accuracy of VGI datasets and their uncertainty demonstrated that the VGI_Editor prototype is capable of validating positional accuracy criteria of integrating VGI and CGDI datasets. In the case of the iPhone_VGI_Before dataset, the percentage of features within the required distance of their true position (assuming GeoBase NRN v.3 layer represents the true position of streets) was increased by 9.23% and satisfies the positional accuracy and its' uncertainty criteria. The GPSMAP_VGI_Before dataset satisfied the positional accuracy and its uncertainty requirements even before validation using the VGI_Editor; nonetheless after such validation dataset accuracy was improved by 4.24%.

The attribute accuracy criterion (Section 2.3) stipulates that attributes should include the following information:

- A name or identifier, e.g., road name for a road; and
- Descriptive information, such as type or function.

Table 4.2 presents attributes of VGI datasets before being validated by users of the VGI_Editor and after it, and Figure 4.6 represents the information on available attributes for the NRN dataset.

Table 4.2 Results of Testing VGI Attribute Accuracy

	VGI Datasets							
Attributues	iPhone_VGI_Before		iPhone_VGI_After		GPSMAP_VGI_Before		GPSMAP_VGI_After	
	Availability	Number of Features	Availability	Number of Features	Availability	Number of Features	Availability	Number of Features
Identifier	✓	33/33	✓	34/34	✓	7/7	✓	8/8
Time Stamp	✓	33/33	P	33/34	✗	0/7	✗	0/8
Length	✓	33/33	✓	34/34	✓	7/7	✓	8/8
User Uploaded	✓	33/33	✓	34/34	✓	7/7	✓	8/8
Editor	✗	0/33	P	1/34	✗	0/7	✗	0/8
Road Name	✗	0/33	P	5/34	✗	0/7	P	1/8
Source of Evaluation	✗	0/33	P	6/34	✗	0/7	P	1/8
Speed Limit	✗	0/33	P	1/34	✗	0/7	✗	0/8
Stop Sign	✗	0/33	✗	0/34	✗	0/7	P	2/8
Bus Routes	✗	0/33	✗	0/34	✗	0/7	P	1/8

Note: "P" means that the provided information is partial.

Province / Territory	Available attributes						
	Basic Attributes	Street & Place Names	Block Face Addressing	Road Jurisdiction	Speed Restriction	Closing Period	Traffic Direction
British Columbia	✓	✓	✓				
Alberta	✓	✓	✓				
Saskatchewan	✓	✓			P		✓
Manitoba	✓	✓	✓				
Ontario	✓	✓	✓				
Quebec	✓	✓	✓				
New Brunswick	✓	✓	✓				
Prince Edward Island	✓	✓	✓	P	P	P	
Nova Scotia	✓	✓	✓				
Newfoundland and Labrador	✓						
Yukon	✓	✓	✓				
Northwest Territories	✓	✓	✓				
Nunavut	✓	✓	✓				

Note: "P" means that the provided information is partial.

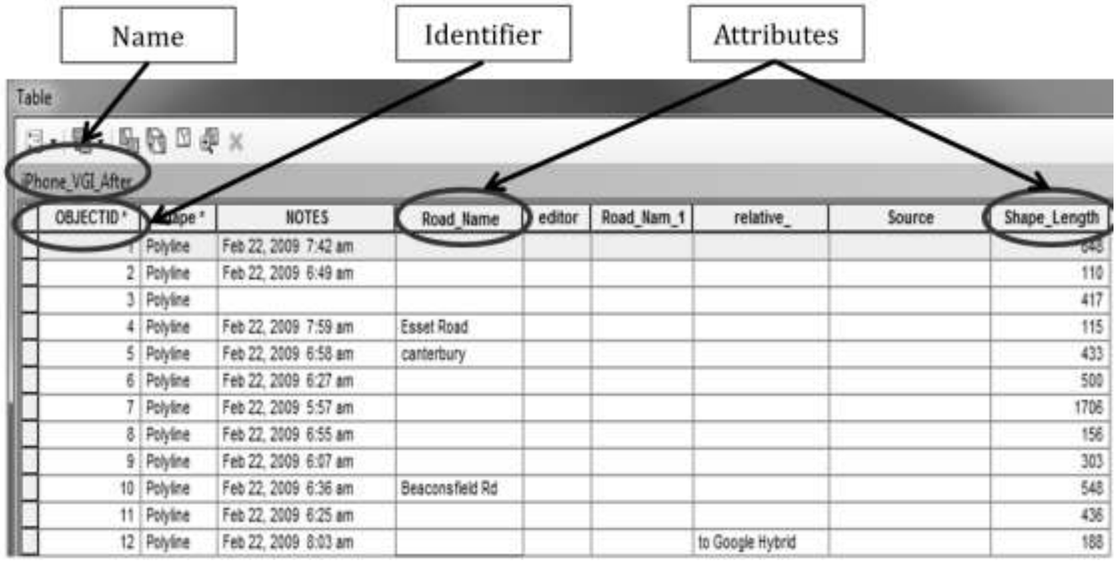
Figure 4.6 Information on Available Attributes for NRN Dataset (from [GeoBase, 2012b])

As shown in Table 4.2 and Figure 4.6, both VGI datasets iPhone_VGI_After and GSPMAP_VGI_After only partially satisfied the attribute accuracy requirements for VGI to be integrated with CGDI datasets. In case of iPhone_VGI_After dataset, only 5 out of 34 features collected had street name attributes provided, while GPSMAP_VGI_After dataset had attributes only for 1 feature out of 8. Both datasets had some basic attributes such as *Source of Evaluation* provided, in case of iPhone_VGI_After 6 out of 34 features and GPSMAPVGI_After 1 out 8 features had such attributes. And none of VGI datasets provided *Block Face Addressing* attribute information.

The uncertainty criterion declares that the maximum acceptable proportion of errors of classification for attributes should be no more than 5%. The uncertainty criterion for VGI attributes was not tested; only identification of VGI attributes was undertaken, because attribute values of NRN and VGI datasets were different, and as such their comparison was not possible. As a result, the VGI_Editor prototype satisfied the attribute accuracy and its uncertainty requirement only partially, as only some attributes were provided for VGI.

4.2.2 Data Structure and Metadata

A data structure requirement states that data should have entities with a name to distinguish them from other entities or features, such as a road or bus stop, and the entities should comprise of: geometric representation, descriptive representation, an identifier and metadata. When exporting, the VGI_Editor prototype renders VGI into a shapefile format which allows for VGI to comply with the data criterion providing a name, identifier, geometric representation and descriptive information for features. All VGI datasets have a geometric representation and data structure of GPSMAP VGI dataset as shown in Figure 4.7 below.



	OBJECTID *	Shape *	NOTES	Road_Name	editor	Road_Nam_1	relative_	Source	Shape_Length
	1	Polyline	Feb 22, 2009 7:42 am						[illegible]
	2	Polyline	Feb 22, 2009 6:49 am						110
	3	Polyline							417
	4	Polyline	Feb 22, 2009 7:59 am	Esset Road					115
	5	Polyline	Feb 22, 2009 6:58 am	canterbury					433
	6	Polyline	Feb 22, 2009 6:27 am						500
	7	Polyline	Feb 22, 2009 5:57 am						1706
	8	Polyline	Feb 22, 2009 6:55 am						156
	9	Polyline	Feb 22, 2009 6:07 am						303
	10	Polyline	Feb 22, 2009 6:36 am	Beaconsfield Rd					548
	11	Polyline	Feb 22, 2009 6:25 am						436
	12	Polyline	Feb 22, 2009 8:03 am				to Google Hybrid		188

Figure 4.7 Data Structure of GPSMAP VGI Dataset

The metadata requirement, described in Section 2.3, for a dataset or a feature stipulates that, at minimum, it should include the source of the data and its accuracy. Metadata information for both VGI datasets is represented through the means of tags in the VGI_Editor prototype and is shown in Figures 4.6.

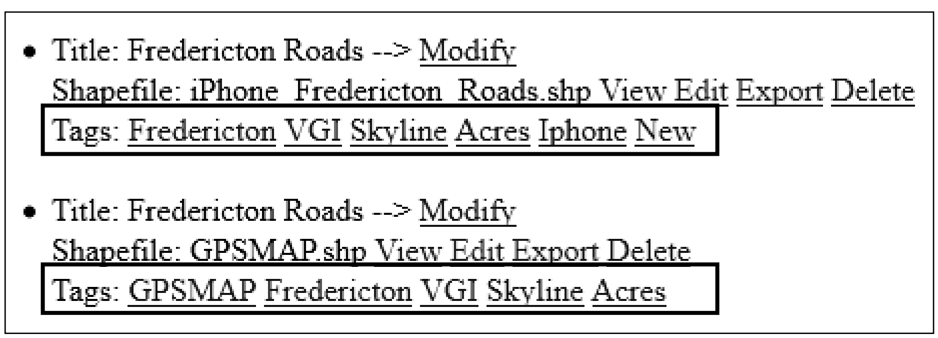


Figure 4.8 VGI Datasets Metadata Information

The criterion for metadata was not entirely compliant because metadata information for both VGI datasets provided information only on the source of the data, and not on its accuracy. The original accuracy of VGI, before being evaluated by users of the

VGI_Editor, could be estimated based on the ‘average’ accuracy of the source of such data. For example in case of iPhone using Assisted GPS technology the accuracy should be less than 10m [Swedberg, 1999; Dao et al., 2002; Barnes, 2003; Sabone, 2009] and positional accuracy for GPSMAP handheld GPS should be better than 10 metres 95 percent of the time [Garmin, 2012]. Accuracy of VGI datasets after being evaluated by users of the VGI_Editor prototype can be estimated based on the following conditions:

- Accuracy of baselayers used for evaluating VGI;
- Results of using BOM for evaluating accuracy of VGI;
- Comments and ratings provided by users of the VGI_Editor prototype; and
- Local knowledge of the area by a user.

Both VGI datasets satisfied the data structure requirement for using VGI as a source of updates for CGDI datasets, but complied with metadata criteria only partially. iPhone_VGI_After and GPSMAP_VGI_After datasets provided information on the source of data, however failed to render the accurate information on their quality, despite providing information that can be used to estimate the accuracy of such VGI datasets. As a result metadata criterion for integrating VGI with CGDI datasets has been met only partially.

4.3 System Evaluation

At the beginning of this research, a set of objectives was defined to help integrate VGI with CGDI datasets. Subsection 4.3.1 discusses how sufficient these requirements have been met by the VGI_Editor prototype.

4.3.1 Research Objectives

The following research objectives defined in Chapter 1 were accomplished to varying degrees in this research:

- *To design and deploy a system (prototype) that is capable of validating VGI to comply with CGDI* positional accuracy *standards in particular:* This research identified accuracy standards to which VGI has to comply in order to be used to update CGDI datasets and the importance of VGI quality evaluation and legal liability concerns in such integration process (Chapter 2). Research subsequently related such evaluation of VGI through design and deployment of the VGI_Editor prototype (Chapter 3) that demonstrated satisfaction of positional accuracy, its uncertainty and data structure in full, and partial satisfaction of attribute accuracy, its uncertainty and metadata standards (Chapter 4) for integration of VGI with CGDI datasets.

- *To discuss and incorporate, where possible, in the prototype some of the legal liability concerns surrounding integration of VGI with CGDI datasets:* Chapter 2 of this research discussed general concepts of legal liability and identified potential sources of legal liability when dealing with VGI. Chapter 2 concluded with a set of RMTs aimed at controlling potential legal liability and Chapter 3 discussed legal liability considerations and functions within the VGI_Editor prototype.

4.4 Summary

This chapter discussed validation performance of VGI against CGDI datasets using the VGI_Editor prototype. The prototype proved to be capable of satisfying positional accuracy and its uncertainty and data structure requirements, while satisfying only partially requirements for attribute accuracy and its uncertainty, metadata and legal liability requirements. The chapter also discussed evaluation of the VGI_Editor based on research objectives. The evaluation showed that some requirements have been met, while others only partially. The next chapter summarizes this research with conclusions and recommendations.

5. Conclusions and Recommendations

This chapter concludes the preliminary development of the VGI_Editor prototype and its implementation. The chapter highlights limitations and lessons learned, as well as conclusions and recommendations for the future research.

5.1 Conclusions

This research was set out to examine whether VGI could be validated to correspond to authoritative, in particular CGDI, datasets accuracy requirements. This research designed a prototype called the VGI_Editor, main function of which is validation of VGI to comply with CGDI accuracy requirements.

This thesis identified that for VGI to be integrated with other datasets of CGDI, it must correspond not only to data quality standards, but also to legal rules and regulations including, for example, intellectual property, privacy, and legal liability. A main contribution of this research was discussion of Risk Management Techniques necessary to limit/reduce potential legal liability issues arising from integration of VGI and CGDI datasets within the VGI_Editor prototype.

Objectives of this research have been achieved only partially, as only requirements towards positional accuracy and its uncertainty, as well as data structure have been met by the VGI. Such requirements as attribute accuracy and its uncertainty, metadata and legal liability have been met only partially, and as such further investigation of different methods to validate and integrate VGI with authoritative datasets is required to improve the prototype. Research outcomes highly depend on the hypothesis made in this research that the GeoBase NRN v.3 layer represents the true position of streets, which might not be the case.

5.2 Legal Issues

This thesis examined legal liability issues under Canadian common law. Specifically the following were discussed:

- Liability in Tort, arising from integration of VGI and CGDI, highly depends on all aspects of VGI data quality, including: positional, attribute accuracy and their uncertainty as well as metadata and data structure. In order to limit such potential liability for CGDI such uncertainties with data quality of VGI should be resolved.
- Liabilities in tort for VGI most likely arise in organizations that incorporate VGI into their datasets, and not to volunteers of VGI.
- To avoid liability, an information provider must be able to show that steps were taken to ensure the accuracy of VGI.

5.3 Research Limitations

The limitations of this research include the following:

- Research covers only the general liability principles concerning VGI and liability in contract and tort. Legal analysis provided in this research should be used to inform and raise the awareness around this issue, and not to provide a comprehensive legal review;
- The methods used to validate VGI did not include a scientific background research and testing;
- Only source of VGI used was a vector data collected using GPS technology, while there are a number of other sources of VGI.

5.4 System Limitations

The main system limitations of the VGI_Editor include:

- There was no prototype design before the development of the VGI_Editor prototype. It is important to define a user base and their system requirements and later test the VGI_Editor for user defined requirements;
- Absence of BOM validation technique within the VGI_Editor prototype, due to time constraints of this research;

- Need for an administrator to initiate validation of VGI using a BOM validation method;
- Need of internet access to use the VGI_Editor prototype;
- Absence of other forms of VGI representation, other than vector format, within the VGI_Editor prototype, such as: pictures, videos and audio;
- Voting system described in Subsection 3.4.6.1 is related to shapefiles only and not on individual features or attributes. Consequently, if one VGI shapefile contains more than one object or more than one attribute then it is not possible to vote for or against them. It is only possible to vote for or against the complete shapefile.

5.5 Recommendations

This research recommends the following:

1. Continue extensive research in the field of crowdsourcing and authoritative approaches towards VGI validation, and their integration within the design of the VGI_Editor prototype until all CGDI requirements towards data quality are met by VGI;
2. Continue extensive research on potential legal liability issues arising from integration of VGI with authoritative datasets. Such research should also be

accompanied by studies on other legal topics, such as: privacy and intellectual property rights. The research should also look beyond the common law;

3. Initiate research on legal liability of CGDI using VGI not only for physical injury but also for economical loss of users;
4. Implement the VGI_Editor prototype: Create a pilot project to determine if it would be feasible for CGDI to update their datasets using VGI validated through the VGI_Editor prototype;
5. Research on and development of tools for actual integration of validated VGI, through the VGI_Editor prototype, with CGDI datasets;
6. Initiate research on examination and comparison of the VGI_Editor functionality with that of the OpenStreetMap [2012] and TomTom [2012] projects to determine development priorities for subsequent versions of the prototype development.

References

About.com. (2012). *"What is a 'tag'? -- What is a 'tag cloud'?"* Web Trends. Retrieved 09/21, 2012, from http://webtrends.about.com/od/glossary/g/tag_def.htm

Adams, T. (2009). *"There is no alternative to OpenLayers ...?"* Presentation at FOSS4G 2009 Conference. Retrieved 09/12, 2012, from Sydney, Australia

Apache. (2012). Retrieved 05/12, 2012, from http://www.apache.org/

Ather, A. (2009). *A Quality Analysis of OpenStreetMap Data.* (Unpublished MEng). University College London, London, UK.

Ball, M. (2010). *"What's the distinction between crowdsourcing, volunteered geographic information, and authoritative data?"* v1 Magazine. Retrieved 10/02, 2011, from http://www.vector1media.com/dialog/perspectives/16068-whats-the-distinction-between-crowdsourcing-volunteered-geographic-information-and-authoritative-data.html

Barnes, S. J. (2003). "Location-Based Services." *E-Service Journal,* Vol. 2(3), pp.59-70.

Basiouka, S. (2009). *Evaluation of the OpenStreetMap quality.* (Unpublished MScE). University College London, London, UK.

Beaumont, C. (2009). *"Wikipedia ends unrestricted editing of articles."* The Telegraph. Retrieved 05/20, 2011, from http://www.telegraph.co.uk/technology/wikipedia/6088833/Wikipedia-ends-unrestricted-editing-of-articles.html

Bédard, Y., Devillers, R., Gervais, M., and Jeansoulin, R. (2004). "Towards Multidimensional User Manuals for Geospatial Datasets: Legal Issues and their Considerations into the Design of a Technological Solution." *Third International Symposium on Spatial Data Quality (ISSDQ),* Vol. 28b, pp.183-195.

Bédard, Y., Gervais, M. and Levesque, M. -. (2013). *"All papers."* University of Laval. Retrieved 03/15, 2013, from http://yvanbedard.scg.ulaval.ca/?page_id=2288&lang=en

Berendt, B., and Hanser, C. (2007). "Tags are not Metadata, but 'Just More Content' - to Some People." *International Conference on Weblogs and Social Media, ,* pp.2012.

Bertrand, N., and Brown, L. (2006). *"Risk Management: A Guide for Nonprofit and Charitable Organizations."* Imagine Canada. Retrieved 08/10, 2012, from http://www.efls.ca/webresources/Imagine_Canada_risk_management_for_non-profits.pdf

Bishr, M., and Kuhn, W. (2007). "Geospatial information bottom-up: A matter of trust and semantics." *The European Information Society,* Vol. Springer, pp.365-387.

Bishr, M., and Mantelas, L. (2008). "A trust and reputation model for filtering and classifying knowledge about urban growth." *GeoJournal,* Vol. 72(3-4), pp.229-237.

Black, H. C. (1979). *Black's Law Dictionary*. St. Paul:. Minnesota: West Publishing Co.

Brown, M. (2001, 17 July). Maps led firefighters up paths with no exits. *The Sydney Morning Herald*, pp. 1-2.

Budhathoki, N. R., Bruce, B., and Nedovic-Budic, Z. (2008). "Reconceptualizing the role of the user of spatial data infrastructure." *GeoJournal,* Vol. 72(3-4), pp.149-160.

Castelein, W., Grus, L., Crompvoets, J., and Bregt, A. (2010). "A characterization of Volunteered Geographic Information." *13th AGILE International Conference on Geographic Information Science,* , pp.1-10.

Chandler, J. A., and Levitt, K. (2011). "Spatial Data Quality: The duty to warn users of risks associated with using spatial data." *Alberta Law Review,* Vol. 49(1), pp.79-106.

Cho, G. (2005). *Geographic information science: Mastering the legal issues*. Australia:. John Wiley & Sons Ltd.

Chrisman, N. R. (1983). "The role of quality information in the long-term functioning of a geographic information system." *Proc.6th International Symposium on Automated Cartography,* Vol. 1, pp.303-312.

Chrisman, N. R. (1991). "The error component in spatial data." *Geographical Information Systems,* Vol. 1, pp.165-174.

Clinton, B. (2000). *"Improving the Civilian Global Positioning System (GPS)."* Office of Science and Technology Policy, Executive Office of the President. Retrieved 06/05, 2011, from http://clinton4.nara.gov/WH/EOP/OSTP/html/0053_4.html

Club TomTom. (2007). *"Get to know MapShare™."* The official blog for TomTom in North America. Retrieved 05/15, 2011, from http://www.clubtomtom.com/general/get-to-know-tomtom-mapshare™

Coleman, D. J., Georgiadou, Y., and Labonte, J. (2009). "Volunteered Geographic Information: the nature and motivation of producers." *International Journal of Spatial Data Infrastructures Research*, Vol. 4, pp.332-358.

Coleman, D. J., Sabone, B., and Nkhwanana, N. (2010). "Volunteering Geographic Information to Authoritative Databases: Linking Contributor Motivations to Program Effectiveness." *Geomatica*, Vol. 64(1), pp.383-396.

Consumer Protection Act, c 30 (2002).

SCR 537 (SCC 2001).

Coote, A., and Rackham, L. (2008). Neogeographic data quality–is it an issue. Paper presented at the *Annual Conference of the Association for Geographic Information, AGI*,

Dao, D., Rizos, C., and Wang, J. (2002). "Location-Based Services: Technical and Business Issues." *GPS Solutions*, Vol. 6(3), pp.169-178.

Delavar, M. R., and Devillers, R. (2010). "Spatial Data Quality: From Process to Decisions." *Transactions in GIS*, Vol. 14(4), pp.379-386.

Devillers, R., and Jeansoulin, R. (2006). *Fundamentals of spatial data quality* (1st ed.). Wiley Online Library.

Digitalpreservation. (2011). *"ESRI Shapefile."* Sustainability of Digital Formats. Retrieved 09/17, 2012, from http://www.digitalpreservation.gov/formats/fdd/fdd000280.shtml

DocForge. (2012). *"Web application framework."* An Open Wiki for Software Developers. Retrieved 09/27, 2012, from http://docforge.com/wiki/Web_application_framework

A.C. 562 (H.L. 1932).

Economist. (1999). *"Power to the people."* Retrieved 05/30, 2011, from http://www.economist.com/node/187034?story_id=E1_TRSDPQ

Edgell, D. F. (2000). *Product Liability Law in Canada*. Markham: Butterworths.

Elwood, S. (2008). "Volunteered geographic information: future research direction motivated by critical, participatory and feminist GIS." *GeoJournal*, Vol. 72(3), pp.173-183.

Estellés-Arolas, E., and González-Ladrón-De-Guevara, F. (2012). "Towards an integrated crowdsourcing definition." *Journal of Information Science,* Vol. 38(2), pp.189-200. Retrieved from http://www.crowdsourcing-blog.org/wp-content/uploads/2012/02/Towards-an-integrated-crowdsourcing-definition-Estell%C3%A9s-Gonz%C3%A1lez.pdf

FGDC. (2012). *"Geospatial Metadata."* Federal Geographic Data Committee. Retrieved 03/05, 2012, from http://www.fgdc.gov/metadata

Flanagin, A. J., and Metzger, M. J. (2008). "The credibility of volunteered geographic information." *GeoJournal,* Vol. 72(3-4), pp.137-148.

Free Dictionary. (2011). *"Tort Law."* Retrieved 11/13, 2011, from http://legal-dictionary.thefreedictionary.com/Tort+Law

Garmin. (2012). *"GPSMAP 76CSx."* GSPMAP 76CSx Specifications. Retrieved 09/25, 2012, from http://www8.garmin.com/products/gpsmap76csx/spec.html

Gartner, G., Bennett, D. A., and Morita, T. (2007). "Towards ubiquitous cartography." *Cartography and Geographic Information Science,* Vol. 34(4), pp.247-257.

Gaurce, M. (2010). *"The Different Types of Registration Software."* Ezine@rticles. Retrieved 09/23, 2012, from http://ezinearticles.com/?The-Different-Types-of-Registration-Software&id=5124775

Genovese, E., and Roche, S. (2010). "Potential of VGI as a resource for SDIs in the North/South context." *Geomatica,* Vol. 64(4), pp.439-450.

GeoBase. (2012a). *"GeoBase Unrestricted Use Licence Agreement."* Retrieved 10/12, 2012, from http://www.geobase.ca/geobase/en/licence.jsp;jsessionid=85A5A28C1444CA5FAA17E07B384DC20B

GeoBase. (2012b). *"National Road Network."* Retrieved 10/14, 2012, from http://www.geobase.ca/geobase/en/metadata.do;jsessionid=57CDDA2DB381077B486132F73DED57AA?id=C0DB1B2E-BDAF-6998-8B8E-E569E5D39D6B

GeoConnections. (2012a). *"GeoConnections."* Retrieved 03/10, 2012, from http://www.geoconnections.org/en/index.html

GeoConnections. (2012b). *"What is the Canadian Geospatial Data Infrastructure (CGDI)?"* CGDI online training manual. Retrieved 03/06, 2012, from http://www.geoconnections.org/publications/training_manual/e/02/02_01/02_01_01.htm

Gervais, M. (2003). *La pertinence d'un manuel d'instructions ausein d'une stratégie de gestion du risquejuridique découlant de la four niture de données géographiques numériques.* (Unpublished PhD thesis). Université La- val, Québec, Canada.

Gervais, M. (2006). "On the Importance of External Data Quality in Civil Law." *Fundamentals of Spatial Data Quality, in Devillers & Jeansoulim, ,* pp.283-300.

Gervais, M. (2012). *"All Papers."* University of Laval. Retrieved 03/14, 2013, from http://yvanbedard.scg.ulaval.ca/?page_id=2288&lang=en

Girres, J., and Touya, G. (2010). "Quality Assessment of the French OpenStreetMap Dataset." *Transactions in GIS,* Vol. 14(4), pp.435-459.

Goeman, W., Martinez-Fonte, L., Bellens, R., and Gautama, S. (2005). Automated verification of road network data by VHR satellite images using road statistics. Paper presented at the *ISPRS Hannover Workshop 2005 High Resolution Earth Imaging for Geospatial Information,*

Goodchild, M. F. (2009a). "Geographic information systems and science: Today and tomorrow." *Annals of GIS,* Vol. -(15), pp.3-9.

Goodchild, M. F., and Hunter, G. J. (1997). "A simple positional accuracy measure for linear features." *International Journal of Geographical Information Science,* Vol. 11(3), pp.299-306.

Goodchild, M. F. (2007). "Citizens as sensors: The world of volunteered geography." *GeoJournal,* Vol. 69(4), pp.211-221.

Goodchild, M. F. (2008). "Commentary: Whither VGI?" *GeoJournal,* Vol. 72(3-4), pp.239-244.

Goodchild, M. F. (2009b). "NeoGeography and the nature of geographic expertise." *Journal of Location Based Services,* Vol. 3(2), pp.82-96.

Google. (2012). *"Google Maps JavaScript API v3."* Google Developers. Retrieved 10/05, 2012, from https://developers.google.com/maps/documentation/javascript/maptypes

Gould Carlson, M., Craglia, M., Goodchild, M. F., Annoni, A., Camara, G., Kuhn, W., Mark, D., Masser, I., Maguire, D., Liang, S., and Persons, E. (2008). "Next-generation Digital Earth: a position paper from the Vespucci Initiative for the advancement of geographic information science." *International Journal of Spatial Data Infrastructures Research,* Vol. 3, pp.146-167.

Gouveia, C., Fonseca, A., Camara, A., and Ferreira, F. (2004). "Promoting the use of environmental data collected by concerned citizens through information and communication technologies." *Journal of Environmental Management,* Vol. 71(2), pp.135-154.

Graff, L. (2003, Better safe: Risk management in volunteer programs and community service. *Ontario: Linda Graff & Associates,*

Grossner, K. E., Goodchild, M. F., and Clarke, K. C. (2008). "Defining a digital earth system." *Transactions in GIS,* Vol. 12(1), pp.145-160.

Haklay, M. (2010). "How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets." *Environment and Planning B: Planning and Design,* Vol. 37(4), pp.682-703.

Haklay, M., and Weber, P. (2008). "OpenStreetMap: User-generated street maps." *IEEE Pervasive Computing,* Vol. 7(4), pp.12-18.

Hourieh, A. (2009). *Django 1.0 Web Site Development*. Birmingham, UK:. PACKT Publishing.

ICBCclaiminfo. (2013). *"Prosecuting a tort claim."* ICBCclaiminfo.com. Retrieved 04/03, 2013, from http://www.icbcclaiminfo.com/node/150

Janssen, K. (2010). *"Legal issues of volunteered geographic information."* International Workshop "VGI for SDI". Retrieved 06/03, 2011, from http://www.law.kuleuven.be/icri/presentations/256VGI_legal_Janssen.pdf

Jones, K. (2011). *Communicating perceived geospatial data quality of 3D objects in virtual globes.* (Unpublished MScE). Department of Geography, Memorial University of Newfoundland, St. John's, Nfld., Canada.

Kalantari, M., Olfat, H., & Rajabifard, A. (2010). Automatic spatial metadata enrichment: Reducing metadata creation burden through spatial folksonomies. In A. Rajabifard, J. Crompvoets, M. Kalantari & B. Kok (Eds.), *Spatially enabling society: Research, emerging trends and critical assessment* (, pp.119-129). Belgium: GSDI Association and Leuven University Press.

Kounadi, O. (2009). *Assessing the quality of OpenStreetMap data.* (Unpublished MScE). University College London, London, UK.

Kuhn, W. (2007). Volunteered geographic information and GIScience. Paper presented at the *Workshop on Volunteered Geographic Information,* University of California, Santa Barbara. , pp.86-97.

Larrivée, S., Bédard, Y., Gervais, M. and Roy, T. (2011). *"New Horizons for Spatial Data Quality Research."* 7th International Symposium on Spatial Data Quality. Retrieved 03/15, 2013, from http://yvanbedard.scg.ulaval.ca/wp-content/documents/publications/604.pdf

Lenders, V., Koukoumidis, E., Zhang, P., and Martonosi, M. (2008). Location-based trust for mobile user-generated content: applications, challenges and implementations. Paper presented at the *Proceedings of the 9th Workshop on Mobile Computing Systems and Applications,* New York, USA. , pp.60-64.

Levesque, M. -., Bédard, Y., Gervais, M., and Devillers, R. (2007). Towards a safer use of spatial datacubes: Communicating warnings to users. *Proceedings of 5th International Symposium on Spatial Data Quality,* Enschede, The Netherlands. , pp.8.

Low, R., Burdon, M., Christensen, S., Duncan, W., Barnes, P., and Foo, E. (2010). Protecting the protectors: Legal liabilities from the use of Web 2.0 for Australian disaster response. Paper presented at the *Technology and Society (ISTAS), 2010 IEEE International Symposium On,* , pp.411-418.

Ludwig, I., Voss, A., and Krause-Traudes, M. (2011). "A Comparison of the Street Networks of Navteq and OSM in Germany." *Advancing Geoinformation Science for a Changing World,* , pp.65-84.

Martinez, J. M. (2003). "Liability and volunteer organizations: A survey of the law." *Nonprofit Management and Leadership,* Vol. 14(2), pp.151-169.

Maué, P., and Schade, S. (2008). Quality Of Geographic Information Patchworks. Paper presented at the *11thAGILE International Conference on Geographic Information Science,* University of Girona, Spain. , pp.1-8.

McCurley, S., and Lynch, R. (1996). "Volunteer Management: Mobilizing All the Resources of the Community." *Downers Grove, Ill.: Heritage Arts,*

McDougall, K. (2009). The potential of citizen volunteered spatial information for building SDI. Paper presented at the *Proceedings of GSDI 11,* Rotterdam, Netherlands.

Mondzech, J., and Sester, M. (2011). "Quality analysis of openstreetmap data based on application needs." *Cartographica,* Vol. 46(2), pp.115-125.

Nkhwanana, N. (2009). *Assessing Credibility of VGI Contributors and Trust in their Contributions.* (Unpublished MScE). Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada.

Obermeyer, N. (2007). *"Thoughts on volunteered (geo) slavery."* Retrieved 05/23, 2012, from http://www.ncgia.ucsb. edu/projects/vgi/participants.html

Onsrud, H. J. (2010). *"Legal Interoperability in Support of Spatially Enabling Society."* Retrieved 05/28, 2011, from http://www.gsdi.org/gsdiconf/gsdi12/papers/907.pdf

Open Geospatial Consortium. (2012). *"About OGC."* OGC. Retrieved 08/30, 2012, from http://www.opengeospatial.org/ogc

OpenStreetMap. (2011a). Retrieved 05/25, 2011, from http://www.openstreetmap.org/stats/data_stats.html

OpenStreetMap. (2011b). *"The Free Wiki World Map."* Retrieved 05/25, 2011, from http://www.openstreetmap.org/

OpenStreetMap. (2012). *"Beginners' guide."* Wikipedia. Retrieved 07/30, 2012, from http://wiki.openstreetmap.org/wiki/Beginners%27_guide

Osborne, P. (2007). *Law of Torts, The. Essentials of Canadian Law Series* (3rd ed.). Irwin Law Inc.

Overton, D., and Abbott, J. (2007). *"Introduction to neo-geography: Embrace the principles of Web 2.0 and use your population to create tomorrows mapping data."* Ordnance Survey. Retrieved 05/24, 2012,

Poser, K., and Dransch, D. (2010). "Volunteered Geographic Information for Disaster Management with Application to Rapid Flood Damage Estimation." *Geomatica*, Vol. 64(1), pp.89-98.

PostgreSQL. (2011). *"The world's most advanced open source database."* Retrieved 06/10, 2011, from http://www.postgresql.org/about/

PostgreSQL. (2012). *"PostgreSQL: The world's most advanced open source database. Advantages."* Retrieved 09/10, 2012, from http://www.postgresql.org/about/advantages/

Rajabifard, A., Kalantari, M., & Binns, A. (2009). SDI and metadata entry and updating tools. In B. Loenen, J. W. J. Besemer & J. A. Zevenbergen (Eds.), *SDI convergence: Research, emerging trends, and critical assessment* (- ed., , pp.121-136). Delft, Netherlands: Netherlands Geodetic Commission.

Rak, A., Coleman, D., & Nichols, S. (2012). Legal liability concerns surrounding volunteered geographic information applicable to Canada. In A. Rajabifard, & D. Coleman (Eds.), *Spatially enabling government, industry and citizens* (, pp.125-143). USA: GSDI Association Press. Retrieved from http://www.gsdi.org/gsdiconf/gsdi13/papers/256.pdf

Rivix. (2012). *"ESRI Shapefile Support."* Retrieved 10/17, 2012, from http://www.rivix.com/shapefiles.htm

Rossmeissl, H. J. (1989). "The spatial data transfer standard: a progress report." *GIS/LIS '89.Proc.Annual Conference',* Vol. 2, pp.699-706.

Rouse, L. J., Bergeron, S. J., and Harris, T. M. (2007). "Participating in the geospatial web: collaborative mapping, social networks and participatory GIS." *The Geospatial Web,* , pp.153-158.

Sabone, B. (2009). *Assessing Alternative Technologies for Use of Volunteered Geographic Information in Authoritative Databases.* (Unpublished MScE). Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada. (Technical Report No. 269)

Safritt, D., Merrill, M., and Nester, N. (1995). "A High Stakes Affair: Managing Risks in Volunteer Programs." *Columbus: Section of Communication and Technology, Ohio State University Extension,*

Scassa, T. (2012). *"Problems with authority: the role of law a barrier to authoritative VGI."* Geographic Information Science and Technology. Retrieved 02/12, 2013, from http://www.gsdi.org/gsdiconf/gsdi12/papers/905.pdf

Scassa, T. (2013). "Legal Issues with volunteered geographic information." *The Canadian Geographer,* Vol. 57(1), pp.1-10.

Sea Farm Canada Inc. V. Denton, C881677 1 (BCSC 1991).

Seeger, C. J. (2008). "The role of facilitated volunteered geographic information in the landscape planning and site design process." *GeoJournal,* Vol. 72(3-4), pp.199-213.

Sieber, R. (2007). *"Geoweb for social change."* Retrieved 05/26, 2011, from http://www.ncgia.ucsb.edu/projects/vgi/supp.html

Sui, D. Z. (2008). "The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS." *Computers, Environment and Urban Systems,* Vol. 32(1), pp.1-5.

Surowiecki, J. (2005). In - (Ed.), *The Wisdom of Crowds* (- ed.). USA:. Doubleday, Anchor.

Swedberg, G. (1999). "Ericsson's Mobile Location Solution." *Ericsson Review,* Vol. -(4), pp.214-221.

Tienaah, T. (2011). *Design and Implementation of a Coastal Collaborative GIS to Support Sea Level Rise and Storm Surge Adaptation Strategies.* (Unpublished MScE). Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada.

Tienaah, T., Rak, A., & Coleman, D. (2013). *An examination and critical comparison of alternative maintenance models for the Nova Scotia digital topographic database.* (Contract Report of 2-year consulting study undertaken for the GeoNova Program Office, Service Nova Scotia and Municipal Relations, Province of Nova Scotia.

TomTom. (2012). *"TomTom - At the Heart of the Journey."* Retrieved 07/31, 2012, from http://www.tomtom.com/en_ca/

Tremper, C. R., and Kostin, G. (1993). "No Surprises: Controlling Risks in Volunteer Programs." *Nonprofit Risk Management Center,*

Turner, J. A. (2007). *"Introduction to Neogeography."* Sebastapol, CA: O'Reilly Media. Retrieved 05/25, 2011, from http://books.google.ca/books?hl=en&lr=&id=oHgDv4feV-8C&oi=fnd&pg=PA2&dq=neogeography&ots=wXt2WFYXeV&sig=QzH-RzLHCs65l9KFRUxPy8WQ-2Q#PPA1,M1

USGS. (1999). *"National Map Accuracy Standards."* United States Geological Survey. Retrieved 02/15, 2012, from http://nationalmap.gov/standards/nmas.html

Wang, R. Y. (1996). "Beyond accuracy: What data quality means to data consumers." *Journal of Management Information Systems,* Vol. 12(4), pp.5-34.

Webopedia. (2012). *"URL."* Webopedia. Retrieved 09/23, 2012, from http://www.webopedia.com/TERM/U/URL.html

Westra, E. (2010). *Python Geospatial Development*. Birmningham - Mumbai:. Packt Publishing.

WikiBooks. (2013). *"Canadian Tort Law/Negligence."* Retrieved 04/02, 2013, from http://en.wikibooks.org/wiki/Canadian_Tort_Law/Negligence

Wikimapia. (2011a). Retrieved 05/25, 2011, from http://wikimapia.org/

Wikimapia. (2011b). *"New Year 2011."* Retrieved 06/10, 2011, from http://blog.wikimapia.org/

Wikipedia. (2008). *"Geobase/Announcement."* OpenStreetMap. Retrieved 07/29, 2012, from http://wiki.openstreetmap.org/wiki/Geobase/Announcement

Wikipedia. (2012). *"Web application framework."* Wikipedia. Retrieved 10/05, 2012, from http://en.wikipedia.org/wiki/Web_application_framework#cite_note-0

Williamson, I. A., Rajabifard, A., & Feeney, M. -. F. (Eds.). (2003). *Developing spatial data infrastructures: From concept to reality* (- ed.). London, UK: Taylor & Francis.

Zielstra, D., and Zipf, A. (2010). A comparative study of proprietary geodata and volunteered geographic information for Germany. Paper presented at the *13th AGILE International Conference on Geographic Information Science,* Guimaraes, Portugal.

Zulfiqar, N. (2008). *A Study of the Quality of OpenStreetMap.org maps: A comparison of OSM data and Ordnance Survey Data.* (Unpublished MEng). University College London, London, UK.

Curriculum Vitae

Candidate's full name:	Andriy Rak
Universities attended:	Lviv National Politechnic University, BE and MScE degrees at the Faculty of Advanced Geodesy and Land Surveying, June 2005.
	Ivan Franko National University of Lviv, LLB and LLM degrees at the Law Faculty, June 2005.
Publications:	Rak, A., Coleman, D., & Nichols, S. (2012). *Legal liability concerns surrounding volunteered geographic information applicable to Canada*. In A. Rajabifard, & D. Coleman (Eds.), Spatially enabling government, industry and citizens (pp. 125-143).
	Tienaah, T., Rak, A., & Coleman, D. (2013). *An examination and critical comparison of alternative maintenance models for the Nova Scotia digital topographic database*. Contract Report of 2-year consulting study undertaken for the GeoNova Program Office, Service Nova Scotia and Municipal Relations, Province of Nova Scotia.
Conference Presentations:	*Legal Liability Concerns Surrounding Volunteered Geographic Information Applicable to Canada*, Session: GEOIDE contributions to SDI research, at GSDI World Conference 2012, 14-27 May, Québec City, QC, Canada.