GIMA Geographical Information Management and Applications

Crowdsourcing and school children?

A quality assessment on VGI collect by children

Case-study Cadastre

Master Thesis

Elisa van Bergen

kadaster

Universiteit Utrecht

TU Delft

Wageningen UR

ITC University of Twente

This page is intentionally left blank

Source figure cover page: School visit at pilot group (2017). Children working with the Web viewer on the computer.

Crowdsourcing and school children?

A quality assessment on VGI collect by children

Elisa van Bergen
Student number: 5697522
e.vanbergen@student.utwente.nl
Masterthesis

Master's programme: GIMA
Place of research: University of Twente - ITC
1ste Supervisor: Frank Ostermann
External supervisor Cadastre: Magdalena Grus
Professor: Menno-Jan Kraak
Time period of thesis: September 2016 – July 2017
Date: June 4th 2017

Preface

In order to complete my master GIMA, students are expected to complete a thesis. The subject of this thesis is to identify a method to analyze the quality of VGI collected by children.

I would like to thank everybody who participated to this project and had a share in its completion. First of all, I would like to thank my supervisor Frank Ostermann for helping me out in trying to understand what VGI is, and giving useful recommendations and advice to structure the thesis. Second, I would like to thank Magdalena Grus from the Dutch Cadastre with which I had interesting discussions about the project. She could support me in gaining information about the possibilities of VGI at the Dutch Cadastre, product information and so on. She also joined me on a school visit which was really nice.

I would also like to thanks my interviewees from which I gained useful information for the thesis: Martin Kraan (Cadastre), Jasper Hogerwerf (Cadastre), and Ana-Maria Olteanu-Raimond (French IGN). As well as Tom Vijlbrief (Cadastre) who supported me in making adjustments within the web viewer.

Furthermore, I would like to thank Tom Moekotte (GIMA) and Stéphane Pulles (GIMA) for joining me on school visits and giving me mental support at the University during coffee breaks.

Last but not least, I would like to thank my family and friends for their support.

Utrecht, June 2017

Abstract: (1) Motivation and societal relevance: In recent years, the advantages of volunteered geographic information (VGI) has attracted the attention of national mapping agencies (NMAs). The Dutch Cadastre has held a pilot-study in 2016 to gain information on the point of interest(POIs) collected by schoolchildren. However, as within all VGI projects uncertainty exists about the quality of data. The task is to identify methods and techniques so that the VGI collected by schoolchildren can be adequately evaluated, since these methods are currently missing. (2) Methods: the methodology exists out of two approaches (i) the methodology to determine the quality assessment method within a case-study, and (ii) a further exploration of this quality assessment method by the use of the datasets collected within the pilot-study. (3) Results: this research shows that the proposed approaches are complementary in assessing data quality. (4) conclusion: The quality assessment methods developed and used within this research are a first step towards a quality assessment method. Based on the results, more research is necessary to discover generic workflows or best practices that could be used to assess the VGI collected by schoolchildren.

Table of contents

List of figures

List of tables

List of abbreviations

BRT	topographical key register
ISO	International Organization for Standardization
OSM	OpenStreetMap
POI	points of interest
RMS	RootMeanSquare
VGI	Volunteered geographic information

CHAPTER 1

Introduction

Due to technological advances, accurate maps have become more and more important in the daily lives of citizens. Google maps for example is used to find information about facilities or activities, while navigation systems as TomTom are developed to navigate towards these places. Another familiar example is the Pokémon Go application released in 2016. Within this application, users see Pokémon's on a map and by walking towards them using GPS, they can try to catch the Pokémon. There are many more examples of applications in which maps are integrated. An advantage is that the applications do not only provide information, but they can also be used to collect information.

Within the 'passive' collection of data the user is providing data without being necessarily aware of it. Often this concerns data stored within social media as geotags attached to pictures or videos. Another example is within the Pokémon go application in which the movement of the gaming activities can be collected (Fast and Rinner, 2014). The 'Missing Maps' event of the Red Cross is an example of the 'active' data collection in which citizen's voluntary gather spatial data to help out the Humanitarian organisation.

The voluntary collection of spatial data by citizens is referred to as volunteered geographic information (VGI) (Goodchild, 2007). What is interesting about VGI is that it has the potential to give competitive advantages to datasets. Because the collection of data is done by volunteers, it provides the opportunity to involve their local knowledge. This can provide specific information, which can be used to enrich, complete or update datasets, or it can be a source for creating new datasets or products (Antoniou and Skopeliti, 2015). Additionally, this collection of data creates the potential to lower data collection costs.

In recent years, the advantages of VGI has attracted the attention of national mapping agencies (NMAs). There have been pan-European meetings such as the *'ICT COST Action IC1203'* and the 'Crowdsourcing and National Mapping Workshops' (EuroSDR, 2017) to address the possibilities of VGI, but certainly also the difficulties. One of the main issues is the uncertainty regarding the quality of VGI, which is in general a result of the fact that VGI is mainly produced by non-experts in varying contexts. This character of data is the opposite of the structured and expert driven context of an NMA (Elwood, Goodchild and Sui, 2012). For this reason, NMAs are looking for reliable processed, techniques and guidelines to assess the quality of VGI. This can assist the NMAs in deciding if VGI can become a part of their spatial data infrastructure (SDI) (Olteanu-Raimond *et al.*, 2016).

The Dutch Cadastre has initiated several VGI pilot studies to fill in this knowledge gap. One of the pilot studies is the '*Crowdsourcing at school!*' project held in 2016. Within this project, children aged 11 and 12 used a Web viewer to collect selected points of interest (POI). Which represent zero-dimensional features referring to specific locations in geographical space, such as a hospital or a rail crossing (Jonietz and Zipf, 2016). The pilot-study was considered a success based on the positive reactions of the children and teachers. However, this does not mean that the pilot-study has been completed. The Cadastre would like to know more about the quality of the VGI collected. Though aforesaid, they do not know which techniques or guidelines they could use in order to assess the quality of the dataset.

Up to now, no studies have been found which discuss the quality assessment of VGI collected by schoolchildren. This makes the opportunities of involving the local knowledge of the schoolchildren to possibly improve datasets under-explored. The purpose of this study is to identify methods and techniques so that the VGI collected by schoolchildren can be adequately evaluated. The main research question of this thesis is:

"How can we assess and ensure the quality of VGI contributed by school children?"

In a broader perspective, the aim of this research is to strengthen the research field concerning the quality evaluation of VGI. Within this research, a case-study is performed to identify which elements of data quality can be used to create a quality assessment method. Within the second part of the research, this method will be used to assess the quality of VGI collected during the pilot-study held in 2016. This helps us identifying what more we can learn regarding the quality assessment of VGI collected by schoolchildren.

1.1 Motivation and societal relevance

VGI can be divided into two main components, i.e. the volunteer and the spatial information (Fritz, See, *et al.*, 2016). Most of the present research explore the latter by assessing the quality of the spatial data collected. Generally, this is done by the use of data collected within OpenStreetMap (OSM) which is a free source map created by volunteers. On the contrary, little is published about the quality of VGI in relation to the volunteers who provide the data. The research of Bartoschek and Keßler, (2012) described the long-term effects of children working with VGI. However, this research does not describe an examination of the quality of the data itself. Therefore, literature about the quality of VGI contributed by schoolchildren has not been found.

This research is an attempt to develop a method to assess VGI collected by children and to make statements about the quality according to this method. The method used within this research is not written as a best practice, but it could possibly help defining best practices in the research field of assessing quality of VGI.

From the societal point of view, the schoolchildren are not only 'used' to collect data for the Cadastre, but they are also taught about the concept of crowdsourcing. According to the Cadastre, it allows the children to practice geographical orientation, and handling of topographical maps. As a beneficial side effect, it could help them to understand the importance of contributing to a map in order to improve it. This relates to the concept of building 'smart cities' in which ICT is used to manage a city's asset (Nam and Pardo, 2011). Making a city 'smart', starts with 'smart' people which in this case, are the schoolchildren.

Background and practical relevance

National mapping agencies as the Dutch Cadastre are important players in the development of national spatial data infrastructures (NSDIs). For a long time, the spatial data has been collected only by NMAs or large commercial companies who had to create spatial data in the geospatial domain. They were the ones who had to model, manage, and update geospatial datasets to ensure sufficient and homogeneous quality throughout a data set (Coleman *et al.*, 2009; Antoniou and Skopeliti, 2015; See *et al.*, 2016). In order to do this, they have developed traditional automated mapping procedures. Yet, a major disadvantage of such procedures is that it are often long-term projects which come along with high costs and restrictive licenses(Elwood, Goodchild and Sui, 2012; Antoniou and Skopeliti, 2015). These disadvantages make it interesting for NMAs to research the possibilities of VGI as an alternative or supplementary source.

The use of VGI is not entirely new within NMAs e.g. the US Geological survey (USGS) allowed citizens to contribute spatial information to the national mapping organization already in 1991(Olteanu-Raimond *et al.*, 2016b). Most of the mapping agencies within Europe have experience in collecting VGI. However, the use of VGI differs as can be seen in Figure 1. The use of VGI at NMAs Especially projects concerning '*change detection'* and the '*reports of alerts'* are integrated within most European NMAs.

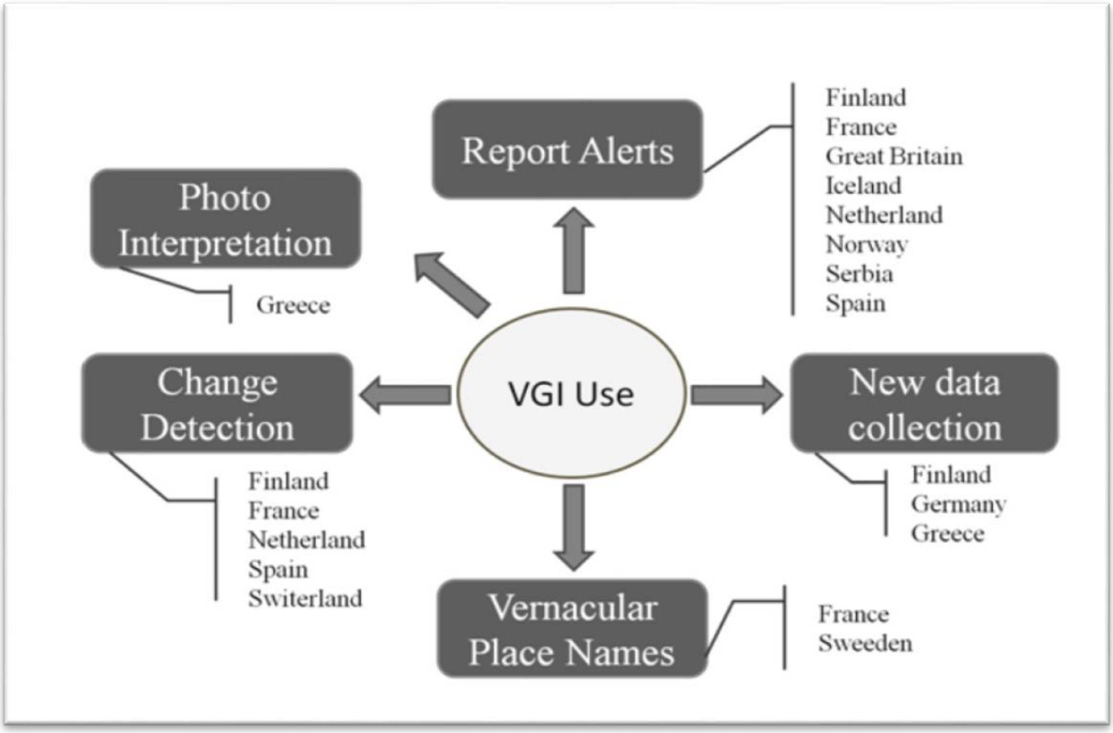


Figure 1. The use of VGI at NMAs (Olteanu-raimond *et al.*, 2016).

The 'Crowdsourcing at school!' pilot-study is a project in which 'new data' is collected. According to figure 1., this type of usage is only known in Finland, Germany and Greece. The Netherlands has not yet integrated the 'new data' collection of VGI (Fonte *et al.*, 2016).

The following workflow (Figure 2.) defined by Olteanu-raimond *et al.*, (2016), is a first approach to visualize the collection of 'new collected' VGI and the integration of this data at an NMA. As can be seen, the collection of VGI is derived from different types of contributors, just as 'schools' which is the central target group of this research.

The workflow works as follows: the contributors must have an interface with which they can make their contributions. These contributions (VGI) have to be assessed on their quality (Q-VGI) which can be done by the contributors as well as the 'experts' at the NMA. The final validation and integration of the Q-VGI can only be done by the NMA. In both the collection of the VGI as the Q-VGI it is proposed to provide feedback to the contributor. The motivation is considered a central part of a project Olteanu-Raimond *et al.*, (2016).

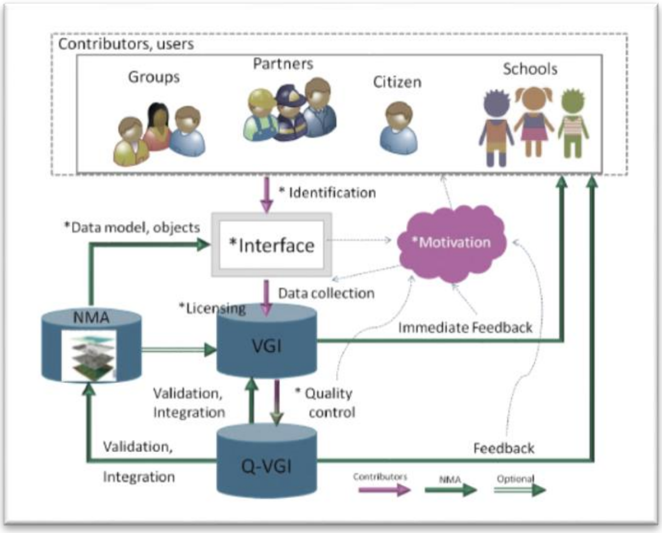


Figure 2. Workflow integrating VGI at an NMA, Olteanu-raimond et al., (2016)

The workflow is a perfect representation on how VGI derived from different contributor groups could be integrated at an NMA. However, since there are currently no reliable processes or guidelines in order to control the quality of data, this workflow adresses the pratical aim of this study: to gain more knowledge on possile techniques and guidelines to assess the quality of VGI collected by children (schools) at a NMA.

1.2 Research Questions

This research will be operationalized by splitting up the main research question into four sub-questions. The research methodology differs per sub-question, which is explained below:

1) ***Which quality measures are used regarding the assessment of VGI?***

   The methods to address this question are the literature study, interviews, a review of the pilot-study and a case-study. This will help to identify which measures are applicable for the assessment of VGI collected by schoolchildren.

2) ***What are the possibilities and issues of VGI projects in the context of the Dutch national mapping agency?***

   To fully understand the 'Crowdsourcing at school project' the context of the project is important. The context of the project will determine conditions set within case-study. The methods to address this question are the literature study and interviews at the Dutch Cadastre.

3) ***What are reliable quality assessment methods for VGI provided by school children?***

   Based on the information in sub-question 1 and 2, two general approaches will be used to explore the quality measures (i) the quality assessment (method) as determined for the case-study, and (ii) an additional evaluation of the quality assessment method by assessing the quality of the datasets collected within the pilot-study.

4) ***What are requirements and strategies to integrate VGI contributed by children into NMA workflows?***

   Based on the result of both approaches mentioned in sub-question 3, conclusions and recommendations will be discussed.

1.3 Research scope

The research scope explains what will be covered in this research, but also what will not be covered:

- Within this research ‘the schoolchildren’ are children aged 11-12 years old attaining a ‘normal’ primary school. No differences e.g. within skills or motivation between the schoolchildren will be made while examining the data. However, this does not mean that the differences are irrelevant for the results of this research.
- This research will focus on the data-oriented assessment of VGI. This means that the contributing crowd, and the web viewer will not be discussed extensively.
- There is little guidance about which quality elements could be used to determine the ‘quality’ of VGI collected by schoolchildren. This makes the decision on which quality elements to use and how to use them project specific. Possibly this research could assist in generating a general workflow or best practice for the quality assessment.
- Currently there is no existing term or concept within scientific literature describing the collection of spatial information by children. Defining the ‘correct’ concept is beyond the scope of this research, but it could be interesting for further research. Out of practical reasons the terms ‘crowdsourcing’ and ‘VGI’ are adopted since these concepts are also used by the Dutch Cadastre.

1.4 Reading Guide

This research is written with the following structure: The upcoming section (Ch. 2) examines the theoretical framework of the research, followed by the Methodology (Chr. 3). What is important to note is that the methodology exists out of two parts (i) the methodology to determine the quality assessment method for the case-study (scientific literature, interviews, and pilot-study assessment) and (ii) the quality assessment method used to perform a quality assessment on the pilot-study. This method is derived from the quality assessment and results of the first part. The result of both quality assessments is discussed in chapter 4. Followed by the discussion (Ch. 5) and the conclusion of the research(ch.6).

Table 1. Overview reading guide

Chapter	Consists of:	Addresses research question:	method	Result
Ch. 1	introduction	-	Literature research	Scientific and practical relevance of the research
Ch. 2	Theoretical framework	*1 and 2*	Literature research	Overview of: - Existing quality measures VGI - Current issues of VGI at NMAs
Ch. 3	methodology	*3*	- Interviews - Pilot-study assessment - assessment case-study results	- Selection of quality measures - Input conditions case-study - Quality assessment method
Ch. 4	results	3	- Field work - results analysis	- Case-study performance and quality assessment results - Quality assessment results pilot-study
Ch. 5	Discussion	1,2,3 and 4	-	- Discussion interpretation results
Ch. 6	Conclusion	1,2,3, and 4	-	- Conclusion research question - Recommendation towards research field and Cadastre

CHAPTER 2

Theoretical background

The first part of this chapter provides background information about the concepts of crowdsourcing and VGI and the quality elements that can be applied. Secondly, an overview of the current issues of VGI within national mapping agencies will be described.

2.1 Key concepts

The concept of VGI has been studied by many researchers using various different terms as 'Neo-geography' (Turner, 2006) or 'user-generated spatial content'(UGS) (Antoniou, Morley and Haklay, 2009). Even though the concepts are often used interchangeably it are not synonyms. The key concepts within this research are 'VGI' and 'crowdsourcing'.

Crowdsourcing was initially introduced by Jeff Howe and Mark Robinson as a business practice in which a business activity is outsourced to the crowd (See *et al.*, 2016). The term implies a low-cost solution by a large number of people which has a value for the set business model. Hereinafter more definitions follow, as the paper of Estellés-Arolas and González-Ladrón-de-Guevara (2012) in which they examined 32 of these definitions. Often these definitions cover the elements concerning the user and the company or organisations proposing the task. As the definition of Goodchild and Li (2012) in which they explain crowdsourcing by the meaning of the terms '*crowd*' and '*outsourcing*', which refers to ''*a number of people who may get to a solution of a problem unable to resolve by an expert or individual*''.

The context of a project defines how the concept of crowdsourcing is interpreted. Within a spatial context Goodchild and Li (2012) describe two interpretations of crowdsourcing. First, the wildfire example, which is explained as a method in which a single observation is strengthened by multiple observations of the same or nearby point. Second is the Linus's law interpretation in which a large group can be expected to edit and correct the errors of others. Both interpretations emphasize the quality assurance of data enabled by the use of crowdsourcing.

VGI is originally coined by Goodchild (2007) in which the concept is defined as: *"the harnessing of tools to create, assemble, and disseminate geographic data provided voluntarily by individuals"*. According to (Elwood, Goodchild and Sui, 2012) VGI is spatial information that is voluntarily made available, with the aim to provide information about the world. Estima *et al.*, (2016) emphasize that VGI gives citizens the possibility to collect and map spatial information which are not determined to be data traditionally mapped by NMAs.

According to (Goodchild, 2007) one of the main differences between VGI and crowdsourcing is that crowdsourcing is recognized as a consensus-producing process provided by a group of individuals while VGI can be collected by one individual i.e. VGI in essence is not determined to do a convergence of others. However, this does not mean that this can still happen while combining the collected data. A well-known example of 'VGI' and 'crowdsourcing' is the OSM aforementioned in the introduction. Contributors can individually add spatial information(VGI) and together with other contributors they 'crowdsource' the OSM.

2.2 Quality measures and indicators of VGI

The 'quality' of VGI is an ambiguous concept since a dataset can be assessed on many aspects. Quality measures often used are from the ISO19157:2013 standard (Haklay *et al.*, 2010; Antoniou, 2011; Keßler, Trame and Kauppinen, 2011; Devillers and Roche, 2013; Jackson *et al.*, 2013). The ISO19157:2013 standard is the most recent international quality standard provided by the 'International Organization for Standardization' (ISO) and is specifically developed to assess the quality of spatial data of authoritative datasets. (ISO, 2013)

The quality measures described in ISO19157:2013 standard are the completeness, logical consistency, positional accuracy, thematic accuracy, and temporal accuracy. These quality measures are also referred to as the 'external' or 'quantitative measures'. They are used to calculate the discrepancy between two datasets, which often concerns authoritative datasets.

Since the character of a VGI dataset differs from the 'structured' dataset of an authoritative dataset, the comparison between these datasets is not always an adequate one. In these cases, the quality indicators: 'purpose', 'usage', and 'lineage', also a part of the ISO19157:2013 standard, can assist. Quality indicators are also referred to as the 'internal' or 'non-quantitative measures' and are mainly used to express the quality overview of the data (Antoniou and Skopeliti, 2015; Fonte *et al.*, 2016; Senaratne *et al.*, 2016). This means that they can be measured without the use of a reference dataset, such as the 'fitness for use' of the project.

Completeness

Completeness refers to the presence(commission) or absence(omission) of features expressed in percentages (Senaratne *et al.*, 2016). Regarding VGI, the quality of the reference dataset is highly important because it influences the value of completeness. For example, VGI that is determined commission because of lacking features in the reference dataset will give biased results.

Another important note is that a dataset may be considered 100% completeness while it does not contain the user needs. This can happen by excluding the needed feature type or by not providing aggregated subtypes when the distinction between those types is necessary. As a consequence, the lack of specifications can make the assessment of completeness a complex process.

Logical consistency

Logical consistency can be defined as "*the degree of adherence to logical rules of data structure, attribution, and relationships*"(Hashemi and Abbaspour, 2015; Fonte *et al.*, 2016). Thera are four main sub-elements considered logical consistency which are: conceptual consistency, domain consistency, format consistency, and topological consistency.

1) Conceptual consistency monitors the adherence to the rules of the conceptual schema by counting the number of features and relationships which violate the conceptual schema;
2) The domain consistency compares the attributes against acceptable domains defined with a certain value by counting the violations compared to the acceptable domains;
3) Format consistency controls the rate of stored data in accordance with the physical data structure by comparing the record structure for all items to field definitions which results in Boolean data;
4) Topological consistency examines the correctness of encoded topological characteristics by counting the number of inconsistencies (INSPIRE, 2013).

The topological consistency in particular, is according to (Hashemi and Abbaspour, 2015) interesting to use within the assessment of VGI. The topological rules can be applied as a real-time check or it can be measured afterwards. For example, a road construction should be placed on or adjacent to a road, while a fire station or a hospital should be located within or adjacent to a building. Smart filters containing these rules, can be used within VGI projects in order to reduce the number of errors.

Accuracy

The accuracy is the degree of closeness between a measurement of quantity and the accepted true value of that quantity(Senaratne *et al.*, 2016). The accuracy can be divided into positional accuracy, temporal accuracy, and thematic accuracy. (i)The positional accuracy is the position of features while comparing them with the reference dataset. Again, the quality of the reference dataset is important to prevent bias. Positional accuracy can be divided in three types of measuring:

1) The absolute or external accuracy is aiming to measure the error distance between the absolute coordinate values of the feature and the reference data. The absolute accuracy can be measured by the RootMeanSquare (RMS) from the error distance.
2) The relative or internal accuracy is measured by the difference of relative coordinate values. This relative accuracy can be measured by the RMS.

3) The gridded position accuracy is the closeness of gridded data position values to values accepted as or being true (Forghani and Delavar, 2014). This measurement can also be calculated by the RMS (INSPIRE, 2013).

Many studies have been done to assess the positional accuracy of VGI regarding OSM (Girres and Touya, 2010; Haklay *et al.*, 2010; Antoniou, 2011; Forghani and Delavar, 2014; Cidália C Fonte *et al.*, 2015). The method of measuring positional accuracy is depending on the type of features used and the purpose of the quality assessment. E.g. for point features the x and y error distance can be measured (Zielstra and Hochmair, 2013) For road intersection the buffer methodology of Goodchild and Hunter (1997) can be used.

(ii) The thematic accuracy implies the correctness of assigned classification to a given entity, with that entity's characteristics and its geographic context. It is indicated by the classification correctness in percentages (Girres and Touya, 2010).

(iii)The temporal accuracy is the quality of the temporal attributes and temporal relationships of features. Which can be divided into 3 types of measurements(Senaratne *et al.*, 2016):

1) The accuracy of time measurement is used to measure the difference between occurrence in the dataset and in the reference data by calculating the RMS on the time differences.
2) The temporal consistency is measured by confirming the temporal order of data which results in Boolean data.
3) Temporal validity is measured by confirming if the date of the acquisition is true or not which also results in Boolean data.

The trade-off of spatial-temporal measurements is that the density and coverage of a VGI dataset over a small time can be limited. Within this measurement, it is necessary to consider the ranges of data appropriate for their purpose and whether cumulative observations are appropriate for their purpose(Fonte *et al.*, 2016).

Purpose and usage (fitness-for-use)

The purpose and usage are indicators used to measure how well the project fits the user requirements i.e. fitness-for-use of the project. The purpose is the intended usage of the dataset i.e. what is the motivation behind the project of the data collection (Senaratne *et al.*, 2016). The usage can be discussed in two parts. First, how well does the data collection fits the user needs of the contributor. And second, the perspective of the project initiator in which question is if the collected data can be used for the purpose of the project.

Lineage

Lineage is the quality indicator that can be used to describe the entire processing history. That is, information about the origin of the data and the way it is collected and evolved (Keßler, Trame and Kauppinen, 2011; Senaratne *et al.*, 2016). Elements that define the lineage are: data description, process description, the date and time, the data source, the responsibility (which identifies the person or parties that performed the process), and the software used.

2.3 Issues of VGI at a National Mapping Agency

According to Olteanu-Raimond *et al.*, (2016b) there are five issues which are considered the main barriers for NMAs to use VGI. The issues are: 'the nature of the crowd and their motivation', 'the sustainability', 'legal issues', 'unemployment fears', and 'data quality'. These barriers will be used as a framework to discuss the issues concerning the integration of VGI into a NMA.

The nature of the crowd and motivation

The volunteers contributing to a VGI project are described by Goodchild, (2007) as 'citizens as sensors'. According to this concept every 'citizen' is able to act as an intelligent sensor by using simple aids as GPS or mobile phones. Within VGI projects as OSM the typical contributor is a well-educated, technology driven male (Mooney, Rehrl and Hochmair, 2013). However, since not every contributor is typical, the 'contributor' is an ambiguous concept. Besides, the 'typical contributors' of OSM can also be sorted on different kinds of characteristics e.g. the varying level of knowledge, technical capabilities or cultural differences (Coleman *et al.*, 2009; Craglia, Ostermann and Spinsanti, 2012; Fast and Rinner, 2014; Antoniou and Skopeliti, 2015). The heterogeneous character of the contributor can create biases within a dataset (Zook and Graham, 2007). This makes it difficult for a NMA to guarantee the quality of the data.

Additionally, the heterogeneous nature of the crowd triggers the interest of the varying motivations of the contributors, such as personal motivations or interests (Antoniou and Skopeliti, 2015; Fritz, Fonte, *et al.*, 2016). The motivations to contribute are just as heterogeneous as the crowd itself which causes another uncertainty on the quality of data. The motivation is also strongly related to the sustainability of a project, as discussed in the following section.

Sustainability

NMAs would like to know how they can keep the contributor motivated (Olteanu-Raimond *et al.*, 2016b). For this reason, different methods are discussed to keep the public engaged. One can think of small payments or the 'gamification' of data collection applications. The latter is the use of game thinking in non-game contexts to engage users to solve problems (Antoniou, V., 2017).

The method chosen is depending on the design of the project. However, the context in which the project is developed is important within this decision. As within the context of a NMA, the method has to be within the boundaries of the regulations. For example, payments should be of 'acceptable' amounts and 'fun games' to collect data should be in line with the 'professional reputation' of the NMA (Olteanu-Raimond *et al.*, 2016b).

Privacy, ethical and legal issues

Public data collections involve many privacy, ethical and legal issues (Mooney *et al.*, 2016). (i) Privacy is the individual's right to have full control of information about themselves, and to decide, when, how, and to what extent the information is shared with others (Oksanen *et al.*, 2015). The data collection of VGI is often not an immediate issue concerning the privacy. However, if the data can be combined with other pieces of information, it could expose private information also referred to as the mosaic effect (GCN, 2017). Currently the guiding principle is to collect as little private data as possible (Mooney *et al.*, 2016). Granell and Ostermann (2015) state there should be a mechanism to ensure privacy if data is gathered. Currently such a mechanism is still to be developed.

(ii) The ethical issues can be discussed from the contributors side as well as the project initiators side(Mooney *et al*., 2016). E.g. the awareness that false reporting can downgrade the value of a VGI project is an ethical approach of the contributors' side. Within disaster management, this could lead to situations in which the map is not fitting the plan of action. From the project initiators side, it is important that the contributors are made aware of the purpose of the project. Additionally, this means that the contributions may not be used for purposes other than stated in the VGI project. Ethical issues from both sides must be taken into account within a VGI project(Mooney *et al.*, 2016). However, not much has been written about how to do this exactly (Granell and Ostermann, 2015).

(iii) Legal issues can be divided into intellectual property and liability. The intellectual property defines the ownerships rights bound to the use of data. For the project initiators, it means that they are bound to rules to disseminate the data. The liability concerns the liability of the contributor itself and under what circumstances the contributions have been done. This is linked with the data quality of the contribution(Mooney *et al*., 2016). The main question concerning the legal issues is: ''who takes the legal responsibility regarding the data?'' There should be a legal protection system in a VGI project to determine how far the responsibility goes. However, again there are no clear rules on how this should be done.

Unemployment concerns

Staff working at NMAs are concerned about the integration of VGI due to the reason that VGI has the potential to substitute a dataset (Olteanu-Raimond *et al.*, (2016b). from the employee perspective, this can suggest the idea that traditional data collection processes could be reduced and jobs could possibly be lost.

Data quality

Data quality is the central issue of this research as discussed in section 1.1. NMAs have the legal mandate to provide spatial data products with a high-quality standard (Olteanu-Raimond *et al.*, 2016b). This makes the unknown quality of VGI one of the main barriers to make use of the VGI.

2.4 Summarizing theoretical background

The quality measures derived from the ISO19157:2013 standard are summarized in table 2. The ISO19157:2013 standard is written to determine the quality of an authoritative dataset. This makes the applicability of the quality elements to assess VGI is questionable. This has to be explored per quality element for the purpose of the project.

Table 2. Summary quality measures ISO19157:2013 standard

Quality measure ISO	Explanation
consistency	Degree of adherence to logical rules of data structure, attribution and relationships. The errors resulting from the lack of it are indicated by the conceptual consistency, domain consistency, format consistency, and topological consistency.
completeness	The absence of data (errors of omission) and presence of excess data (errors of commission) within %
Positional accuracy	Accuracy of position indicated by the absolute or external accuracy, the relative or internal accuracy, or the gridded data position accuracy.
Thematic accuracy	The classification correctness within %
Temporal accuracy	The accuracy of the temporal attributes and temporal relationships of features
Quality indicators ISO	
Purpose	Describes the intended usage of a dataset
Usage	Describes how the dataset has been utilized *(user- and projects initiator perspective)*
Lineage	The source of data, as well as how it was collected and what transformations took place

From the context of the NMAs there are five main barriers to use VGI. Which are: 'the nature of the crowd and their motivation', 'the sustainability', 'legal issues', 'unemployment fears' and 'data quality'. These issues will be used to give a recommendation to the Cadaster within the final part of this research.

CHAPTER 3

Methodology

As foresaid in the reading guide the methodology exists out of two parts (i) the methodology used to determine the quality assessment method for the case-study and (ii) the quality assessment method developed within the case-study applied to the quality assessment of the pilot-study. Prior to the methodology, the following section clarifies why first the case-study is performed and thereafter the assessment of the pilot-study.

The case-study provides the possibility to:

- Interact with the children to get an indication of possible difficulties within the web viewer or recommendations in order to improve it. Additionally, the reaction of the children on the project itself is interesting to explore its potential;
- Choose the data collection method i.e. teaching lesson itself and the web viewer being used. Possible adjustments within both can be made.
- Collect data within one research area from schools that are all placed within this area. This provides the possibility to choose a research area with a high density of POIs, but which is small enough to examine manually all objects within a reference dataset. Additionally, this is interesting for measures as the completeness in which the question can be asked how much of the POIs the schoolchildren can find within the area. Also, the relatively large number of data within the same area is interesting, because this would also be the case if the project would become a part of the teaching curriculum. The potential of the number of contributions can be explored.

The pilot-study:

- Within the case-study, the pilot-study provides the possibility to give an indication on how the case-study can be performed (data collection method, and data collected). This assist the scientific literature and the interviews to determine the conditions of the case-study, resulting the quality assessment method of the case-study;
- Thereafter, the data collected within the pilot-study is used to further explore the quality assessment method determined within the case-study.

3.1 Part 1: Methodology case-study

This section describes the research methods used (interviews and assessment pilot-study) to determine a quality evaluation method for the case-study.

interviews

In total 4 interviews are conducted. The following table (table 3.) shows the name of the expert, their expertise, and the reason of the interview. The general goal of the interviews is to gain knowledge rather than to make conclusion. For this reason, guiding questions were used, instead of a formal topic list, to perform the interview. The interview questions and summarized answers can be found in appendix II.

Table 3. Overview interviewees

Name	(pilot) study	Reason to interview	interview	Date
Martin Kraan	Feedback system	To gain knowledge about the quality assessment of the feedback system (error and report web viewer).	Face-to-face	8-11-2016
Jasper Hogerwerf	Toponyms pilot-study	To gain knowledge about another VGI project in which new data collection is performed.	Face-to-face	8-11-2016
Ana-Maria Olteanu-Raimond (French IGN)	Contribution articles: 'Crowd and community sourced data quality assessment', 'VGI quality control', 'crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information' and many more.	To discuss her ideas about a quality assessment of VGI regarding the context of the case-study.	Skype session Email contact	21-11-2016 28-02-2017
Magdalena Grus (thesis supervisor)	'Feedback system', 'forest paths on the map', 'crowdsource at school project'	To gain knowledge about the crowdsourcing at school project and to discuss ideas about the quality assessment and the necessity of crowdsourcing in general.	Face-to-face	Various during thesis

Summarized: Relevant findings for the case-study based on the interviews:

- The Dutch Cadastre has no experience on the assessment of quality of 'new' collected VGI collected by children. Grus (2016) would like to obtain knowledge about the assessment of quality. Specifically, the research needs to answer how many of the contributions of the schoolchildren are 'correct' or 'incorrect';
- Quality measures are project specific e.g. positional accuracy, thematic accuracy and completeness only makes sense if reference data is available. Topological rules can be added within the contributing process or can be calculated afterwards by the number of violations of the topological scheme Olteanu-Raimond (2017);

- In order to perform data-matching processes VGI can be compared with data sets provided by the Cadastre as the BRT (topographic key register) and the BAG (key register buildings). Additionally, the Cadastre has aerial imagery and 360° photographs to check objects which cannot be retrieved from the reference datasets Grus (2016);
- A quality indicator of the VGI could be the number of contributions for one object-type on the same place (wildfire example) (Hogerwerf, 2016).
- A voting system can be integrated into the Web viewer as an additional quality check, besides the visualisation of contributions of others could serve as a reward. It would be interesting to analyse if this quality indicator can indicate how 'good' children can judge the contributions of others (Grus, 2016).

Pilot-study assessment

This section will discuss the used web viewer, the methods of data collection and the VGI itself. A screenshot of the Cadastre web viewer is shown in the figure 3. Within the Web viewer the children can add 10 different POIs. The selection of objects is determined by the Cadastre based on the request from the emergency services sector to provide more detailed information. Additionally, the objects are chosen based on the theme 'emergency services'. By giving the Web viewer this theme it is aimed

Figure 3. Screenshot Web viewer - pilot-study

to make it understandable for the children that accurate maps are important e.g. if there is a fire, it is important that the fire station knows if the building is a primary school or a hospital.

The objects chosen are: primary schools, high schools, fire stations, police station, hospitals, train station, rail crossing, bus station, gas station, and 'pole'. The name of the object-type 'pole' is ambiguous. What is meant are poles op to waist-height intended to block access of motor vehicles to streets or paths. For an example, see Figure 4.

Figure 4. Screenshot 360° picture 'pole' (Cyclomedia)

The Web viewer works as follows: the children can add their name and password (not obligatory). They can zoom or type in the address of the place they would like to go to. While they are searching, they can switch between a topographical map(BRT) and aerial images. This can help them to recognize the POIs. Once they think they found an object, they can click on the map and specify the object. The objects can be added on the topographical map as well as the aerial images in which no differences are made. They can add as many objects as they would like.

The resolution of the map and aerial imagery is 42cm per pixel with a scale 1:1191. The zoom level is 13(standard levels). The maximum zoom level of the BRT and the corresponding aerial imagery are shown in the screenshots in Figure 5 5. We can assume that this zoom level should give a feasible possibility to place the objects on the intended places.

Figure 5. Zoom level Web viewer- pilot-study

Data collection method

The method is determined by the 'teaching method' and the possibilities within the Web viewer e.g. the zoom level, or the visibility of contributions of others. The methods used within the pilot-study and the feedback written down by the Cadastre is displayed in Table 4.

Table 4. Data collection method - pilot-study

School	Notes feedback form
Wageningen	One child stood at the digital board in front of the class while other children were helping. The teacher was playing a 'control' role on the children. It took time to discuss where to object needed to be placed which eventually resulted in 4 data points. A benefit is that a small amount of data is relatively easy to process. A disadvantage was that not all children were actively involved in the discussion. To get more children involved with the lesson the method was adjusted into working in groups of 2-3 persons
Apeldoorn	There were 2-3 children working on a computer and were split in two groups. The first group could add objects without seeing contributions of others. The second group could add objects while seeing the contributions the first group.
Wolfheze	There were 2-3 children working on a computer and could not see the objects added by other children. It resulted in a high amount of data in which many of the added objects related to the same location. This could be seen as a disadvantage. however, on the other hand, the high number of objects pointing to the same location can be treated as a sort of validation. E.g. the objects mentioned by at least 8 of 10 groups can be named as very highly reliable, 6 - 8 of 10 high reliable, 2-6 of 10 medium reliable (where extra control/validation is needed), less than 2 not reliable and the object can be neglected.
Utrecht	There were 2-3 children working on a computer. All the children could see contributions of the other children while working with the Web viewer. The assumption was that if children could see the contributions of others they are triggered to find other places where objects are not added yet. On the other hand, it can be said that if children already see the contributions of others it is quite easy to add the same object. The assumption cannot be confirmed nor rejected.

Within the pilot-study four primary schools across the Netherlands were visited in which 128 children contributed to the project. Together the children collected 236 objects. In Table 5. the numbers of collected data can be found per school. As can be seen, the numbers of data differ per school. These differences can have several causes. However, within this case big differences are due to the differences between data collection methods e.g. Wageningen versus Utrecht.

Table 5. Data collection - pilot-study

Schools	Location	Number of children	class	Date	Data(objects)
School 1	Wageningen	29	8	26-11-2015	4
School 2	Apeldoorn	40	2 times 7	25-01-2016	88
School 3	Wolfheze	28	8 and 7	27-01-2016	67
School 4	Utrecht	31	8	23-05-2016	77

By visualising the data onto a standard base layer in Arc GIS most of the objects are placed within the surroundings of the school. Only a few contributions are placed way outside of the surrounding of their school. We can assume that this will also be the case within a new data collection.

The meta-data of the point features are slightly different. This is due to the changes made within the web viewer per data collection method. In general, the meta- data contains information about the date and time, object-type, and x and y coordinates. Missing in the metadata is the age of the contributor.

Findings relevant for the case-study based on the pilot-study

- The data collection method (teaching method and web viewer) is relevant for the dataset being collected. From the teaching method perspective, the preference (based on the Cadaster's experience) is that the children first get a classical explanation followed by groupwork (2-3 children). In this way, everybody receives an instruction of the Web viewer and they are explained why it is important to contribute. Additionally, working in 'small' groups triggers the children to contribute. From the Web viewer perspective, it will be interesting to design the web viewer in a way that the children cannot see the contributions of other while contributing themselves Main reason is that children have to actively, and the number of independent contributions from one real world object can be used as an indicator of validation.
- We can assume that the density of potential contributions by the children is correlated with the distance to their home and their school, i.e. we cannot expect contributions on areas that are far from the visited schools.
- The contributions are often placed adjacent to the reference feature to which it is referring to. An example can be seen in figure 5. The hospital and school adjacent to the building are 'correct' but not placed on the building itself.
- The same type of POIs can be used for the case-study since their quality has not been assed (yet).
- Since the web viewer is openly available, it would be useful to integrate a question about age into the web viewer of the case-study. Other information is not necessary.

Quality measures Case-study

The following table (table 6.) lists the quality elements described in section 2 and additional quality indicators based on the interviews and feedback within the pilot-study. The methods deemed suitable (in green) and partly suitable (in orange) are described in more detail in the following section.

Table 6. Quality measures

Quality measure ISO	Measurement method
Positional accuracy	*Not suitable –the VGI consists out of POIs representing different kinds of features. A comparison with reference data representing other features (polygon, line) makes the comparison difficult. As a solution, the centroid of a building could be used as a reference point. However, with large buildings having a complex layout (e.g. hospitals), this could lead to distorted and unreliable results.*
Temporal accuracy	*Not suitable – Since the VGI is collected during a school's curriculum with a specific moment in time, the information is not assumed to be representative.*
Topological consistency	Partly suitable – rules integrated into the web viewer (object buildings on buildings, 'Pole' on road, 'train station' and 'rail crossing' on rail track. Otherwise the validation question ''are you sure you want to place the object here?'' is asked.
Thematic accuracy	The classification of correctness of objects matched on proximity, in comparison with the object-type. This will be calculated by the following formula: $$\frac{Matched\ data\ (thematically)}{Matched\ data\ (proximity)} * 100$$
completeness	The absence of data(omission) and the presence of excess data (commission) This will be calculated by the following formula: $$\frac{Observed\ number\ of\ features\ in\ VGI\ \text{dataset}}{Total\ number\ of\ homologous\ features} * 100$$
Quality indicators ISO	
Purpose	The Cadastre would like to know how the quality assessment can be performed in order to know the ratio 'correct' or 'incorrect'.
Usage	The usage is existing out of two parts, the user needs of the contributors, and the user needs of the projects. The user need of the project is defined by the purpose, which will be calculated by the matched- not matched equation: $$\frac{\text{Matched data (proximity and thematically)}}{\text{Total number of POI in VGI dataset}} * 100$$
Lineage	Description of data collection and necessary transformations
Additional quality indicators	
Multiple contributions	Validation of a POI based on the number of contributions
Voting	Validation based on voting values

Thematic accuracy

Thematic accuracy addresses the question whether a specific data point represents the same feature in the 'real world'. Within the web viewer, children can determine where to place the objects. An aggregated thematic accuracy on a per-feature basis can be measured by calculating the ratio of thematic matches of data-matched to the objects based on proximity. For example, 10 data-matches within proximity of a train station, and 4 of them have the attribute 'rail crossing' while 6 have the attribute 'train station' then the thematic accuracy of this feature will be 60 %. This thematic accuracy is calculated for all reference objects matched to VGI point based on proximity.

Completeness

Within the case-study the same types of POIs (10 in total) will be used within the web viewer. In order to assess completeness, objects of these types have to be present in the reference data. This is the case with the exception of the object-type 'pole'. Therefore, the completeness is calculated on the remaining 9 object-types.

The overall completeness likely depends on the feature density within the study area. I.e. we can expect a study area with a lower density (e.g. Wolfheze) to have a higher overall completeness than a study area with a high density of objects (e.g. Utrecht).

Topological consistency

The topological consistency can be integrated into the web viewer as a second validation question (e.g. ''are you sure you want to place the object here?''. The measure is not used within the quality assessment afterwards since it is noticed that children often place contributions adjacent to the feature. Since this is not necessarily incorrect, it is difficult to define a topological scheme to determine 'violations'. However, if the project would be held on a bigger scale with large amounts of data it would probably be useful to integrate this measure to scale down the number of 'false' contributions.

Purpose and usage

The purpose of the project is to identify techniques and methods to assess the quality of VGI collected by children in order to determine the objects 'correct' and 'incorrect'. This is calculated by matching the VGI with the reference data, based on proximity and object-type. VGI that meets the requirements are considered 'correct', while those that fail any of the matching requirements are considered 'incorrect'. With the object-type 'pole' the data-matching is based on a comparison between 360° photographs provided by the Cadastre and the contributions of the children.

The usage will be discussed in two parts. First, how well does the data collection fits the user needs of the contributor. And, second, can the collected data be used for the purpose of the project to determine the ratio 'correct' or 'incorrect' contributions. To measure this ratio the new equation 'matched- not matched' is developed calculating the matched features based on proximity and object-type

Lineage

The lineage will cover the description of the different school visits and the transformations necessary to measure the differences of quality within the collected data.

Multiple contributions

This quality indicator can be used to obtain an indication of the completeness of the reference dataset i.e. if multiple children place the same object on the same place it may be assumed that the object has to be there in the real world. If the reference dataset does not contain the same information, the dataset will be checked with using external sources (e.g. 360° photographs). If correct updating of the reference data set is required.

Validation based on votes

The possibility to vote on contributions of others can be integrated into the web viewer. These votes could function as an indicator of the quality of data. It will be interesting to analyse the positive and negative votes to identify the quality of data based on these votes. Possibly voting can help to pick out ‘incorrect’ data.

Case-study

The case-study is defined by the research area, the web viewer, and the data collection method.

We chose Hilversum as a study area, because it is a middle-sized municipality with a high density of POIs, but which is small enough to examine manually all objects within the reference dataset.

A couple of adjustment are made within the web viewer see: appendix III. The first validation step is during contributing. Within the Web viewer topological rules are added. If a child places an object somewhere else than written in the topological rules, a question pops up to ask them if they are sure that they want to add the contribution. They can check the box if this is the case, otherwise they can replace the object. Additionally, the web viewer does not show the contribution of others while the children are contributing themselves. The number of independent contributions from one real world object will be used as an indicator of validation. In a second screen the children can see the contributions of others and vote (agree or disagree) This is can be done after they have completed their own contributions. The votes will be accumulated per contribution giving a final voting value, i.e. if 4 children vote negative and 2 positive the value of votes shows minus 2.

We contacted all 34 schools in Hilversum multiple times by email and telephone. This resulted in positive responses of 3 schools from which 2 schools were willing to contribute with two groups, making it a total of 5 groups. The visits were conducted within a short time window. The scheme of the lesson can be found in appendix IV. The data collection is done by 2-3 children having 10 minutes to add objects in the Web viewer, followed by 5 minutes to see and vote on the other contributions.

Before the actual data collection, a pilot session was held at a primary school. This is not used for the quality assessment, but led to small adjustments to increase usability of the web viewer, e.g. the simplification of the URL to start the web viewer. Additionally, the pilot session helped to improve the introductory information and explanation of the exercise to the children.

Quality assessment method Case study

The conceptual model of the quality assessment is visualized in figure 6. The quality assessment starts with defining the purpose of the project followed by the data collection. The topological consistency rules are integrated in the data collection phase.

The collected VGI needs to be pre-processed for the data-matching processes by dividing the dataset per-object. These datasets will be matched with the reference datasets, which also needs to be divided per-object. While comparing the VGI with the reference data, the reference datasets will be checked on its completeness according to the quality indicator 'multiple contributions'. Subsequently, the data 'matched' can be used to measure thematic accuracy and completeness. The 'matched- not- matched' equation will be used to calculate the ratio between both. Additionally, the quality indicator 'votes' is examined by the use of these datasets to analyse the quality of the voting behaviour of the children.

In green are the ISO measures, in grey the lineage of data, and orange the additional quality indicators. The results of the quality assessment of the case-study are described in chapter 4.

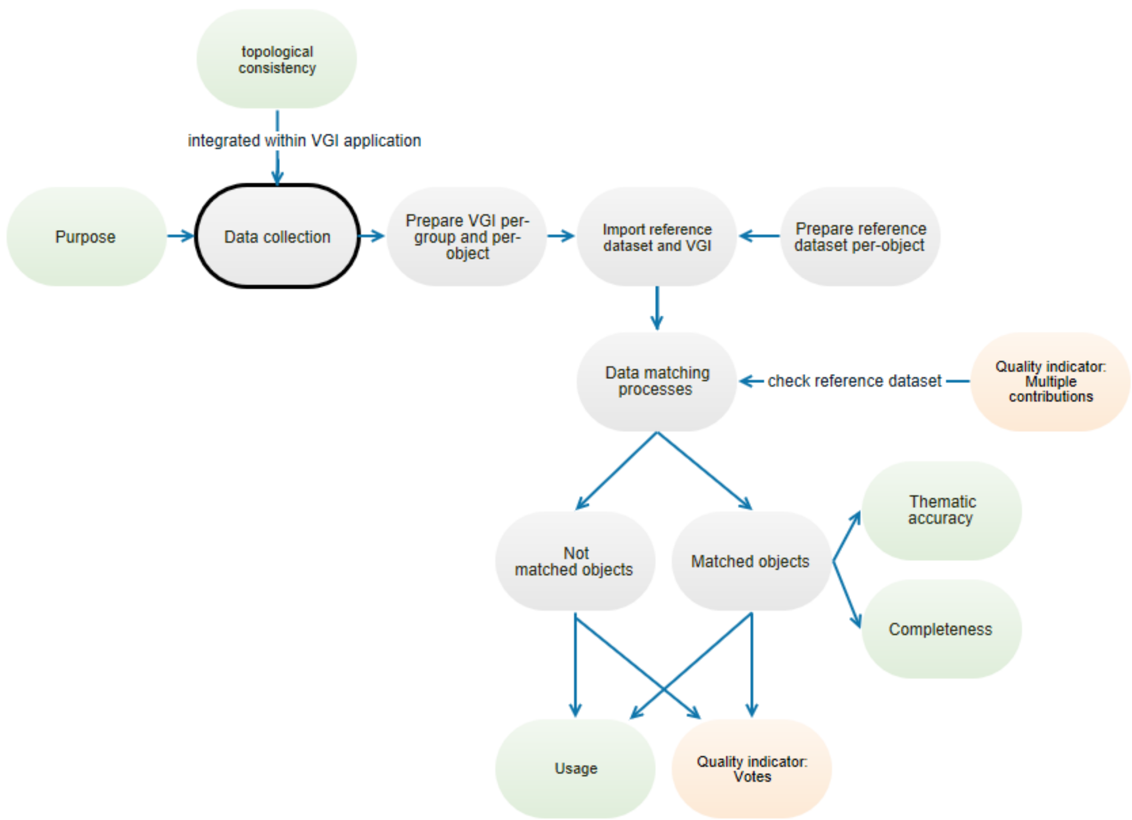


Figure 6. Conceptual model quality assessment (case-study)

3.2 Part 2: Methodology Pilot-study

Within the second research approach the datasets collected within the pilot-study are used to further explore the quality assessment method and to be able to compare the quality of both datasets. However, since not all measures are applicable for the pilot-study a selection of the quality measures is made. The following table (table 7.) lists the quality elements as described in section 2. Based on the case-study results the method deemed suitable (in green) and partly suitable (in orange) are used within the pilot-study assessment. The applicable measures are used in the same way as within the case-study, as described in section

Table 7 Quality measures - Pilot-study

Quality measure ISO	Measurement method
Positional accuracy	*Not suitable – Same as within the case-study*
Temporal accuracy	*Not suitable -- Same as within the case-study*
Topological consistency	*Not suitable – not integrated within the web viewer. Additionally, all data are taken into account, so also 'violations' according to a topological scheme.*
Thematic accuracy	$\frac{Matched\ data\ (thematically)}{Matched\ data\ (proximity)} * 100$
Completeness	*Not suitable – Based on the case-study results the completeness is not an appropriate measure*
Quality indicators ISO	
Purpose	The Cadastre would like to know how the quality assessment can be performed in order to know the ratio 'correct' or 'incorrect'. This will be calculated by the following formula: $\frac{Matched\ data\ (proximity\ and\ thematically)}{Total\ number\ of\ POI\ in\ VGI\ dataset} * 100$
Usage	The usage can be discussed in two parts, how well does the data collection fits the user needs and can the collected data be used to assess the data on objects. determined 'correct' or 'incorrect'.
Lineage	Description of data collection and necessary transformations
Additional quality indicators	
Multiple contributions	Validation of a POI based on the number of contributions
Voting	Not suitable – not integrated within the Web viewer

Quality assessment method Pilot study

The conceptual model of the quality assessment is visualized in figure 7. Compared to the case-study, this quality assessment method does not contain the: topological consistency, completeness, and voting indicator.

Again, it starts with defining the purpose of the project followed by the data collection. The collected VGI needs to be pre-processed by dividing the dataset per-school (differences in methods) and per-object. These datasets will be matched with the reference datasets which are also divided per-object. While comparing the VGI with the reference data, the reference datasets will be checked on its completeness according to the quality indicator 'multiple contributions'. Subsequently, the data 'matched' will be used to measures thematic accuracy. The matched and not- matched data will again be used to calculate the ratio between both.

Deemed in green are the ISO measures, in grey the lineage of data, and orange the additional quality indicator.

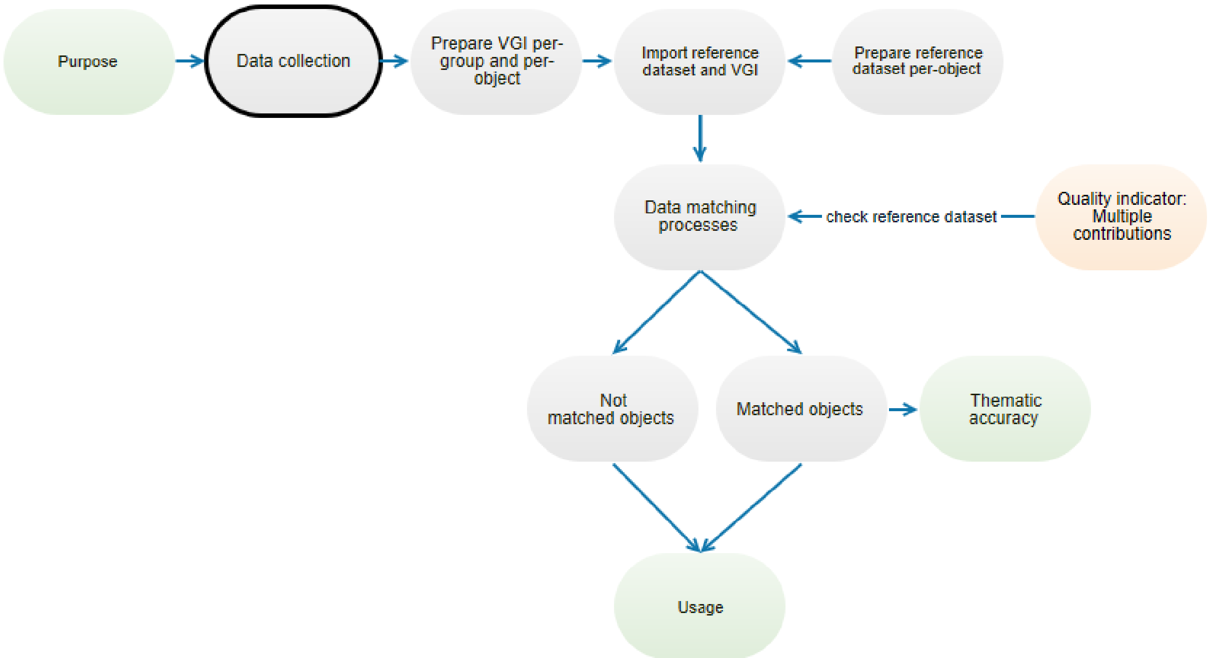


Figure 7. Conceptual model quality assessment (pilot-study)

CHAPTER 4

Results

This following chapter consists the case-study results and the pilot-study results of the quality assessment method.

4.1 Case-study results

The total contributions in Hilversum are visualized in figure 8. As expected the data is somewhat clustered e.g. the JP Minckelerschool in the east with the contributions in red. However, this does not mean that the children do not reach out of the 'border' of their assumed living area. As can be seen in the north(-west), there was no school visit there but still there are multiple contributions.

Due to the small sample size, it is not possible to do statistical tests. However, descriptive analysis is used to discuss the results.

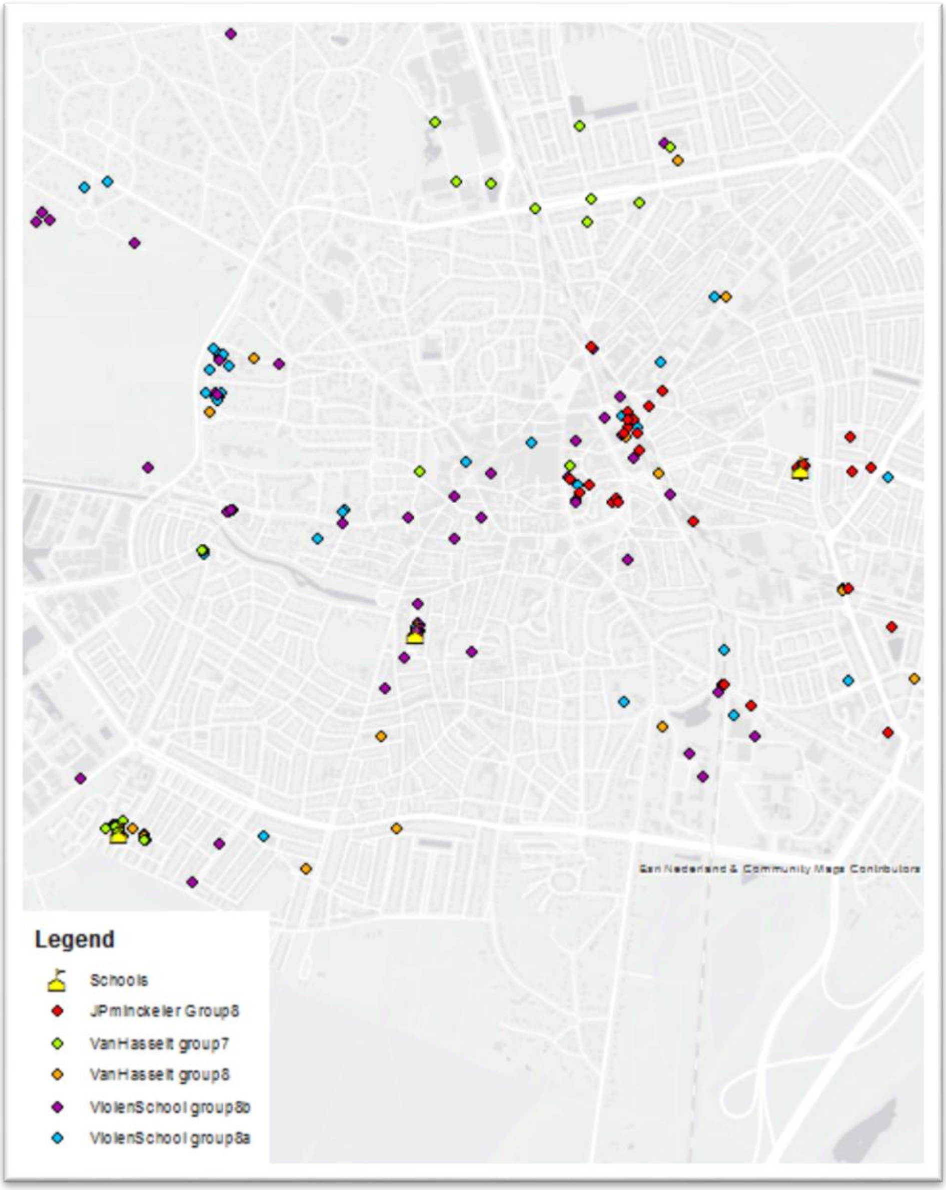


Figure 8. Data collected within Hilversum

Data lineage

Within table 8. the results of the data collection are shown. As expected most of the contributions are placed within the research area with an average of 86.6%.

Table 8. Data collected within case-study

School	Date	Number of children	Amounts of data in total	Data in Hilversum	% data in Municipality of Hilversum
JP Minckeler school - group 8	21-03-2017	23	44	42	95.5
van Hasselt - group 8	27-03-2017	27	33	31	93.9
Van hasselt - group 7	27-03-2017	23	40	32	80
Violenschool - group 8a	12-04-2017	24	76	62	81.6
Violenschool - group 8b	12-04-2017	18	69	60	87
Total	-		*262*	*227*	***86.6***

The VGI is divided into 10 classes representing the different types of objects. This is done per-group, so (5 times 10 classes) and for the dataset as a whole. The resulting datasets are used for the data-matching processes.

data-matching

The data-matching procedures are based on the proximity to the reference object and the object-type. Because both datasets differ from each other, the data-matching procedures caused some difficulties.

Within the case-study the POIs represent a range of 10 different object-types, this makes it challenging to find corresponding reference datasets of the POIs. Additionally, they could be named otherwise e.g. within the reference dataset a ‘gas station’ is labelled ‘industrial function’, but this does not mean that every object labelled ‘industrial function’ is a ‘gas station’. Also, because the POIs represent varying geographic extents it is difficult to determine general rules for all objects. As for example, features can have several disjoint polygons e.g. a school having multiple buildings. Some children might place their contribution on the main building, others in the school yard. Thus, more than one possible placement is correct. In order to tackle these difficulties, a strategy has to be determined. Within this research the VGI is compared to the reference datasets by the use of buffers calculated within Arc GIS. More will be explained in the following sections.

Reference data

To assess the VGI data collected, reference data is necessary. The objects used within the Web viewer are extracted from multiple reference datasets as PDOK, TOP10NL, and the BAG see appendix I. Additional layers provided by ESRI and the UU database are used to cover some thematic layers e.g. primary- and secondary schools.

Except for the polygons labelled 'industrial function' which includes the 'gas station', there was no reference data found on the object type in particular. For this reason, addresses of gas stations are derived from an external source which is validated by the use of 360° photographs. A new layer file of the obtained information is created in Arc GIS. Because 'gas stations' sometimes have buildings and sometimes they have not, a point layer is created to designate the locations. 'Poles' cannot be obtained within external sources. For this reason, only the 'poles' appointed by the children will be controlled and validated by the use of 360° photographs.

The number of POIs and feature type is shown in table 9. The object-type 'train station' is, maybe different than expected, also represented as a point feature. Just as with gas stations this object sometimes has a building as the central train station while others do not (only platforms). Additionally, the train stations building is only a part of the whole central station i.e. the rail tracks at the station are not included. The reference dataset of train stations represents points features which are placed within the 'middle' of a train station. This is the same as for the object-types 'bus station' and 'rail crossing'.

Table 9. POIs and feature types

Points of interest	Number of objects	Feature type
Primary School	34	Polygon
Secondary school	13	Polygon
Gas station	15	Point
Police station	1	Polygon
Hospital	1	Polygon
Train stations	3	Point
Rail crossing	5	Point
Bus station	1	Point
Fire station	1	Polygon
Pole	unknown	Point

Figure 9. POIs visualised at Hilversum

figure 9. on the next page, the POIs within the research area are visualized. As can be seen the object-type 'pole' is not included.

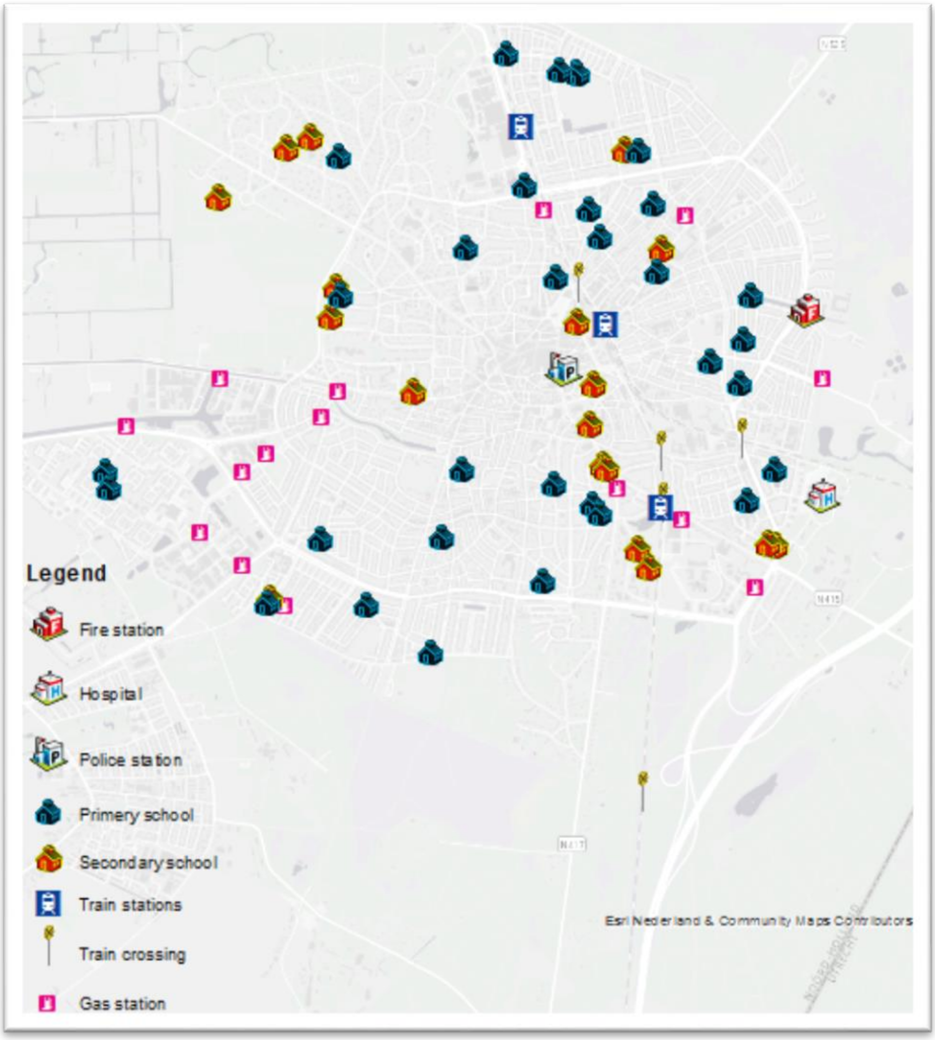
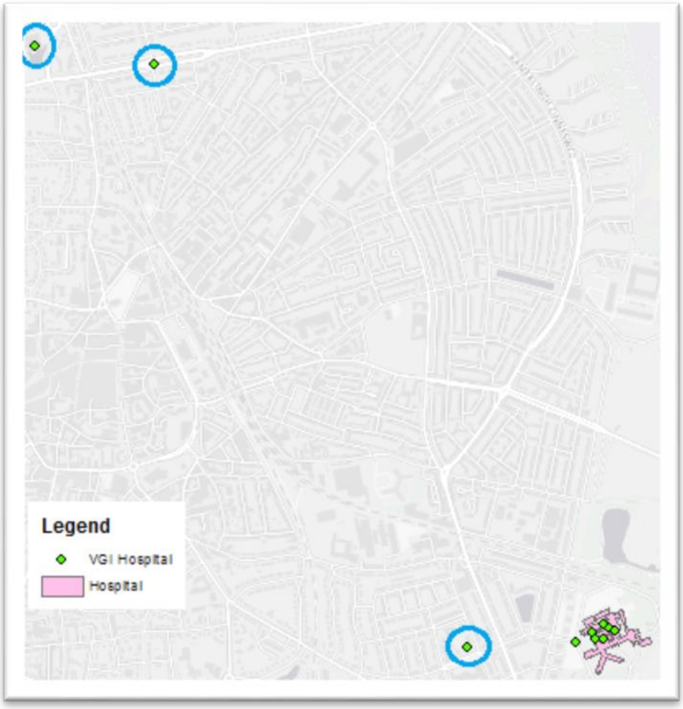


Figure 9. POIs visualised at Hilversum

Defining buffers

By buffering the reference objects, we allow for some level of placement inaccuracy and complex (ambiguous) geometries of the reference objects. We used some experiments and heuristics to determine acceptable buffer distances.

An example of such as experiment is explained by the use of figure 10. Based on the reference data, we can assume that the 3 outer data objects in the left image (circled with blue) are not placed correctly. If we take a closer look (right image) 1 objects is not placed on the building. However, this does not mean it is incorrect. To capture this data point, the nearest distance between the data point and the building is calculated and rounded up.



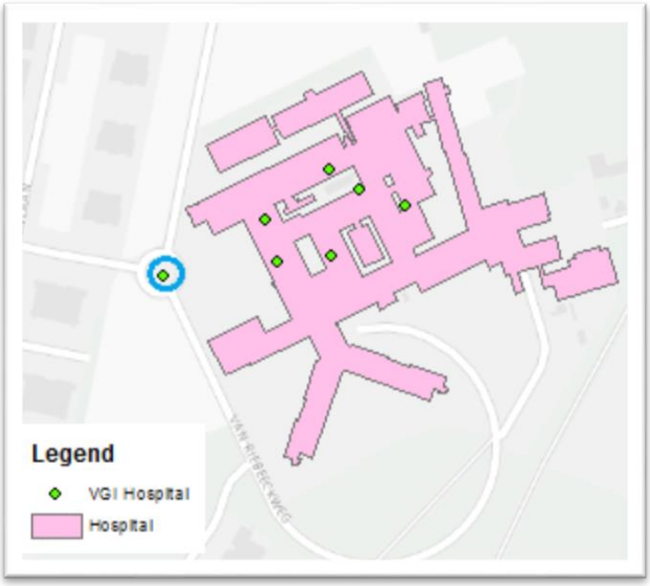


Figure 10. Determine buffers

The buffer calculations are performed per-object which results in 9 datasets. The results are shown in table 10. Subsequently the VGI datasets and the reference datasets are imported to Arc GIS. Depending on the quality measure being calculated the data is compared mostly by using the 'select-by-location' tool.

Table 10. Buffer distance per POI

Points of interest	Buffer distance in meters
Primary School	35
Secondary school	35
Gas station	30
Police station	35
Hospital	35
Train stations	80
Rail crossing	35
Bus station	30
Fire station	35
Pole	-

4.2 quality assessment results (case study)

The quality assessment is based on the quality measures and quality indicators. The measure are completeness and thematic accuracy. The quality indicators are the purpose of the data collection and the intended usage in which the matched- not matched equations is used. Subsequently, the quality indicator 'votes' is examined. The results are shown in appendix V, VI, VII, VIII, and IX. Short descriptions on the results are discussed in the following section.

completeness

During the quality assessment, it became clear that the completeness is not suitable for the quality assessment within this case-study. The given results have to be interpreted as results of editing speed, because the children were subjected to time, rather than giving the completeness of a dataset.

Within the per-group comparison it can be seen that the total percentages differ per school. The Violenschool group 8 and b have relatively high rates, followed by the Minckelerschool. The van Hasselt school group 7 has relatively low rates. Only the schoolchildren at the Minckelerschool have found an object of every object-type. It variates on which objects has not been found per school class, but is seems that 'fire station', 'police station', and 'hospital' are used less. By observing the results of the total dataset, it seems that the higher the number of objects the lower the completeness, however, this seems logical since it is easier to find one hospital instead of 15 gas stations within time limitations.

Thematic accuracy

The Minckeler school has a high rate on thematic accuracy. That means that from the objects found 98.5% of the contributions has been chosen 'correct'. The van Hasselt school group 7 has the lowest thematic accuracy rate of 55.9%.

By evaluating the quality results on the total dataset, the object-type 'gas stations' stands out with a 21 out of 21 correct. The object-type 'train station' on the other hand has a remarkably low percentage. While analysing the data, this is because of the use of 'rail crossing' instead. Surprisingly this does not work the same way the other way around. The lower percentage at the feature type 'bus station' is mostly because children also used this object to point out a bus stop. The percentage thematically correct on the whole data set is 75.7%.

Purpose and usage

Since these quality indicators give a description of the overall dataset, the purpose and usage will be further described within the 'Discussion' chapter (Ch. 5) of this research. However, the results of the matched- not-matched equation, set as a part of the purpose and usage of the project, will be discussed here. The results of the calculations can be found in appendix VII.

The results are based on the ratio of the total collected VGI compared to the VGI determined correct based on proximity and thematically. Again, the Minckelerschool has relatively high match rates. The

Violenschool 8a and 8b (especially b) has relatively low percentages compared to the high rates previously calculated within the quality measures completeness and thematic accuracy. Simultaneously for the van Hasselt school group 7 it works the other way around. They have relatively high rates compared to the measures previously calculated. In total, they have 59.4% of the total VGI contributed 'correct'.

While observing the results of the total dataset the object-type 'gas station' is used relatively often by the children with high match rate of 95.5%. The feature type 'fire station', 'police station', 'hospital' are used less with lower match rates. The object 'bus station' has the lowest match rate. Within this measure the object-type 'pole' is taken into account. Based on this comparison with 360° pictures 50% of the contributions is 'correct'. The average match rate on all data is 70% i.e. 70% of the contributed data is determined correct.

Quality indicator: Votes

During the data collection 1314 votes are given by the children. Compared to the number of contributions added to the map (227) this is almost 6 times more. The results of the votes variate from minus 6 towards a positive value of 8. Within table 11. the number of objects with the averaged assigned number of votes is shown. From the 227 contributed objects, 39 contributions have no value which means that 39 contributions did not get a vote. Simultaneously this means that 82% of the contributions within the dataset did get a vote. 24 Of the contributions have a value of zero which means that the accumulation of negatives and positives votes resulted this value.

Table 11. Voting values

Votes value	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	No value
Number of objects	3	5	2	8	12	32	24	41	24	12	12	5	3	3	2	39

In total 62 contributions got negative values while 102 contributions have averaged positive values. Within figure 11. on the next page, the negatives and positives are visualized. We can see that the contributions with a negative average (in blue) are placed more outside the buffers of the reference dataset then the positive voted contributions (in green). This could mean that the negative votes help to get violations out of a data set.

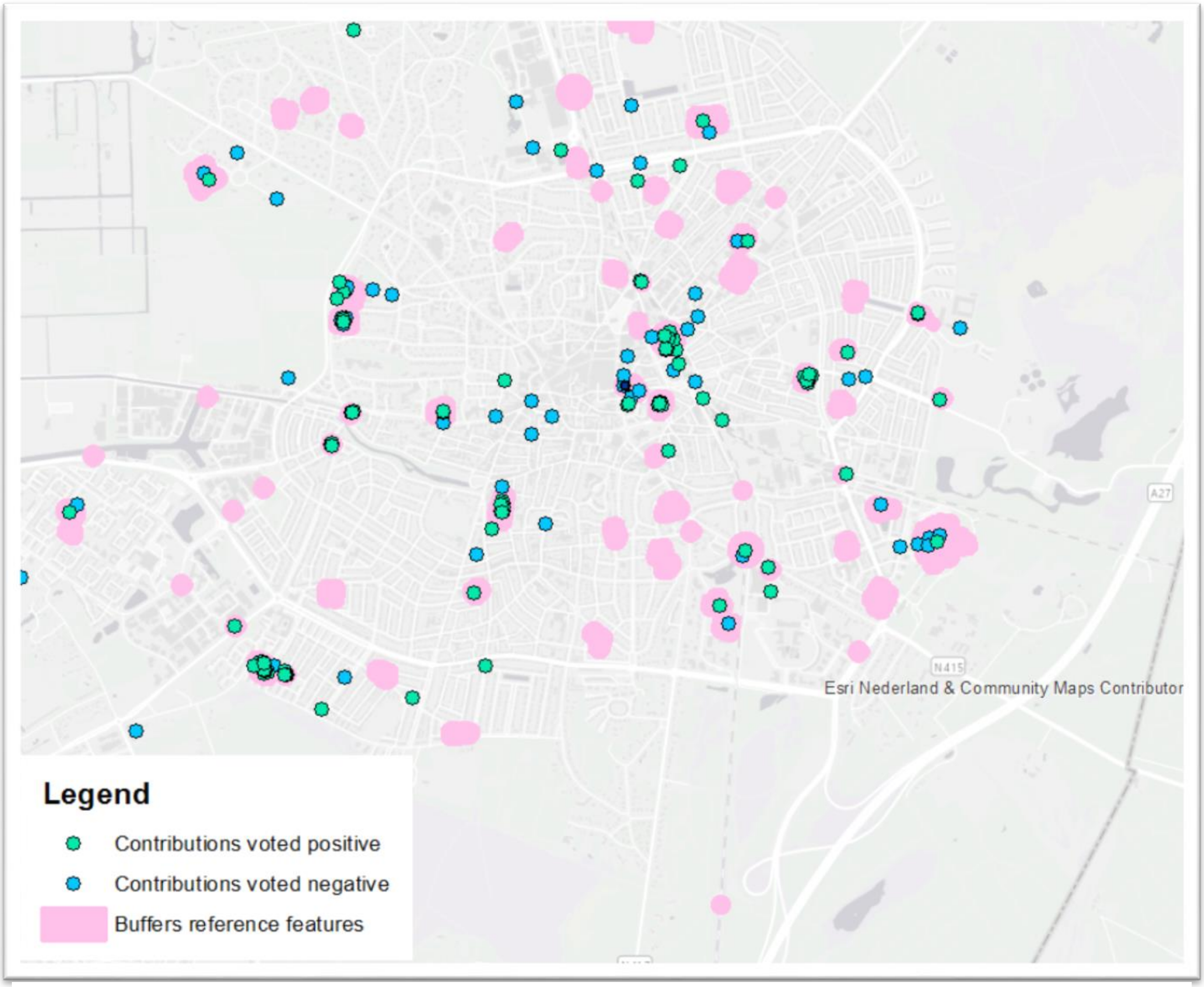


Figure 11. Negative and positive votes visualized

Another interesting observation is visualised in figure 12. It seems like the children found the contribution placed in the middle of the object 'the best'. This could mean that the voting system could be used the reduce the number of contributions for the same object which makes a data analysis easier.

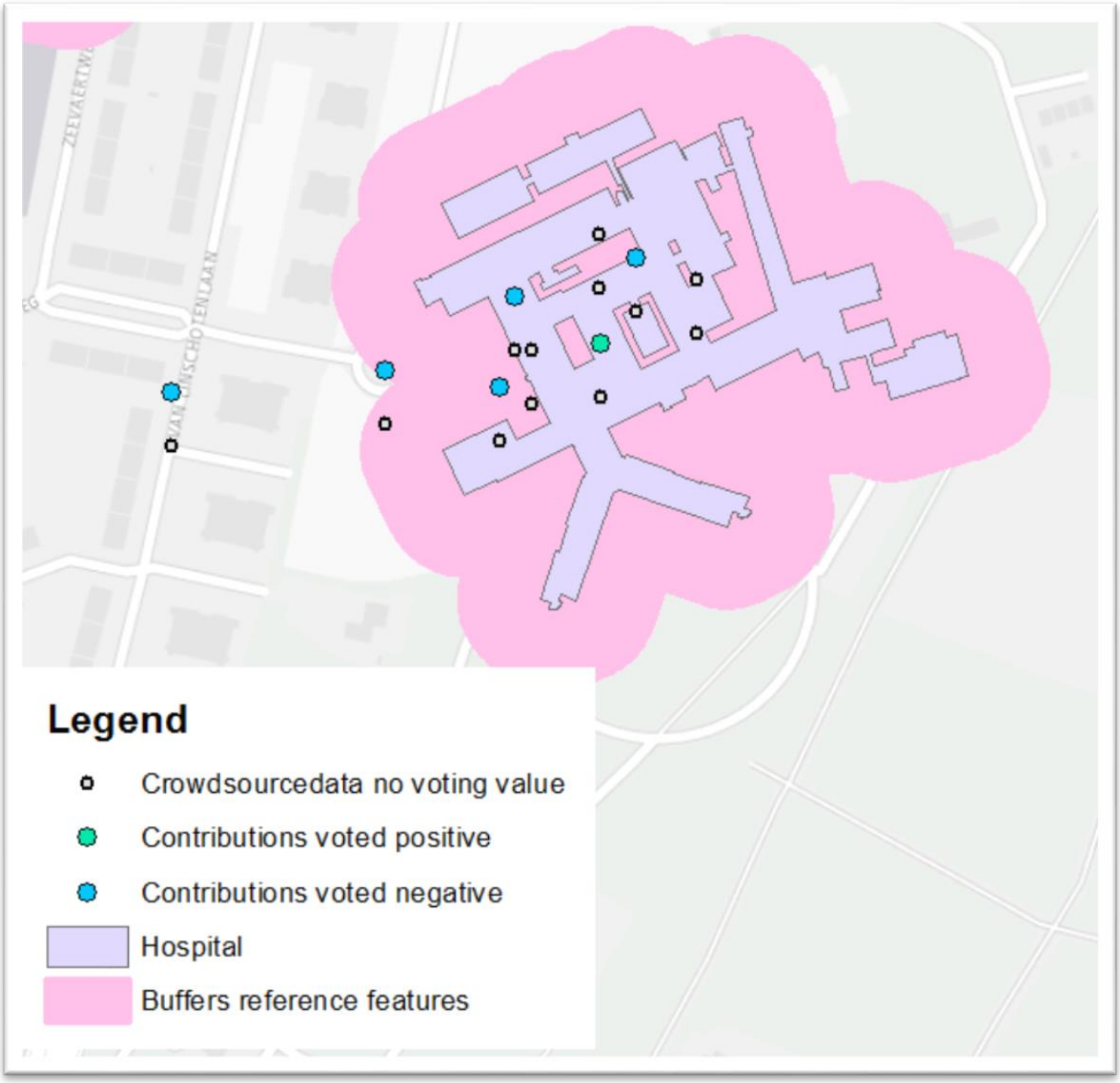


Figure 12. Screenshot zoom-in on hospital

To assess if these observations are correct, a quality assessment is performed in which both the positive as negative votes are assessed. The quality measure used are thematic accuracy, and matched-not matched. Completeness is not used since this measure is not suitable. The results can be found in appendix VIII.

Within the results of the matched- not-match indicator, the 'negative' voted contributions matching to the reference dataset is 45.1%. This means that 54,9% of the negative voted features are not matched and by this determined incorrect. In comparison to the positive votes which have a match of 84.3% which means that 15,7% of the positive voted contributions is determined incorrect. We may assume that the votes based on these results can give an indication on the quality of the data.

4.3 Pilot-study results

The pilot-study assessment is performed on the 4 datasets collected within the pilot-study held in 2016. More information about the methods used during the data collection are previously described in section 3.3.

Data lineage

Within table 12. the results of the data collection are shown. As expected most of the contributions are placed within the research area with an average of 93.2%.

Table 12. Data lineage - Pilot-study

School	Date	Number of children	Number of data in total	Data within municipality	% data in Municipality of Hilversum
Wageningen	26-11-2015	29	4	3	75
Apeldoorn	25-01-2016	40	88	77	87.5
Wolfheze	27-01-2016	28	67	64	95.5
Utrecht	23-05-2016	31	77	76	98.7
Total	-	*78*	*236*	*220*	***93.2***

The VGI is divided into 10 classes representing the different types of objects. This is done per-school, so (4 times 10 classes) and for the dataset as a whole (9 classes). The resulting datasets are used for the data-matching processes.

Data-matching

The same data-matching method is used as described within section 4.1.1 (data-matching), so first the reference data must be defined. Since the pilot-study is done within 4 different municipalities, the reference data is identified for every municipality. The reference data is withdrawn from the dataset as described in 4.1.1(reference data).

Within the dataset of the municipality of Wageningen there are only 3 objects added within the municipality. Only these object-types are analysed to identify which reference objects were necessary within the reference dataset. This saved time in the data preparation for the reference dataset in Wageningen. Within the data-matching procedures buffers are used as described in 4.1 (defining buffers, table 10.).

4.4 Quality assessment results (pilot-study)

The quality assessment is based on the thematic accuracy the results are shown in Appendix IX.

Thematic accuracy

The thematic accuracy of the data collected in Wageningen is 100%. This is a high accuracy rate. However, the low number of data must be taken into account while interpreting this results. The thematic accuracy of the data collected in Wolfheze seems relatively low. While analysing the data, this is mostly due to the object-types 'train station' and 'rail crossing'.

Apeldoorn and Utrecht have relatively high accuracy rates. Except the object-type 'primary school' seems low. While looking at the data, this is because the children used the object-type 'pole' to point out poles standing on the school yard. These contributions are taken into account based on the buffer used(35m) defined in the methodology.

Purpose and usage

The purpose and usage of this research will be discussed within the 'Discussion' chapter of this research. The results of the calculations of the matched- not matched equation set as a part of the purpose can be found in appendix X.

Again, the VGI collected within Wageningen has a high match rate, especially in comparison to the results of the other schools. The match rates on proximity and object-type are relatively low e.g. Apeldoorn 51.1% and Utrecht 68.4%. While analysing the object-types this is mostly caused by the contributions pointing out the object 'secondary schools', 'police stations', and 'Poles'.

CHAPTER 5

Discussion

5.1 General observation case-study

The data collection was successfully done within a short time span. The children seemed to enjoy to work with maps within an interactive way. What was noticed during the lesson is that even though the set-up of the teaching method was the same, every class reacted in a different way. For example, the visits at the JP Minckeler school and the Violenschool were during the morning in which the children were sharp and actively participating. While at the van Hasselt school it was late in the afternoon and warm outside through which the children were a little more distracted. This is understandable, but could clarify differences within the resulting data.

Another noteworthy observation were the differences in skills and knowledge. Some children already heard about the concept of crowdsourcing or could understand it quite quickly, while others needed some extra explanation. Some children found the task easy while others were struggling to recognize objects from above as in comparison with street view.

As a final, it was interesting to observe that even though the importance of contributing was explained, some children were making 'jokes' during contributing i.e. deliberately adding incorrect data. E.g. one child was making contributions on where he wanted the objects to be in the real world.

5.2 Quality assessment Case- and pilot-study

There are various factors influencing the quality of the data, as the teaching method, the user interface of the web viewer, the objects used within the web viewer and the contributing crowd itself. Since there is a limited number of literature available, and there was no example quality assessment method, it is important to reflect on the outcomes which could help within further research concerning this topic.

The quality assessment of the case-study is based on 5 groups with a total of 115 children within the municipality of Hilversum. One method of data collection was used. Even though one method was used the results of the quality measures varied per group. This can have several causes e.g. differences in knowledge, skills, and atmosphere within the group. These differences are not analysed within this research, though it does remind us on the heterogeneous character of the crowd.

Within the pilot-study different methods were used at the different schools. What was noticed in the data analysis is that within methods in which the children could see the contributions of others, multiple contributions could state the same but were 'incorrect'. This could insist that children copy each other if they see the contributions of others.

The method of Wageningen resulted only 4 data points, but they were all identified 'correct'. During the data analysis, this saved quite some processing time. This could be an advantage for the Cadastre. It will depend on what the Cadastre requires from the children to find if this method is sufficient enough. For example, if they are asked to point out the police station not that many contributions are necessary to cover this object. However, objects as 'primary schools' or 'gas stations' in general have a high density of objects. It will be difficult to cover these object-type with such a small number of contributions.

Within the quality assessment of the case-study as well as the pilot-study a combination of quality measures and quality indicators is used to define a suitable quality assessment method and calculate results. The results will be discussed in the following section

Quality measures

For the (external) quality measures, data-matching processes were necessary. The data-matching method is project specific. This means that the results of this research must be analysed in context of the object-types used and the buffer ranges determined. The quality measures are discussed in the following sections.

Logical consistency

Within the case-study this measure is used as an extra validation method during the contribution of objects. The children were told that they had to tick the box if they were sure that their contribution was 'correct'. The children were not told about these rules, because it could influence their choice on placing object e.g. also correct contributions not existing in the reference dataset.

Looking back at this quality measure it was not successful. The question was more disturbing than helpful for the children. Simultaneously it does not fulfil the purpose of the project i.e. the children are checked if they placed the object 'correct' according to the reference dataset, rather than assuming that the contribution can be complementary to the reference dataset. In a future project, I would remove this measure from the web viewer.

The completeness

Aforementioned the completeness does not seem an appropriate measure within the quality assessment. For this reason, the quality measure was not calculated within the pilot-study. This measure was rather an indication of editing speed, than that it said something about the children completing the POIs in the reference dataset i.e. the results within this research represent the editing speed of 115 children all having 10 minutes in total to contribute.

Within future research the children can be asked to perform the task at home without this time restriction. Simultaneously, just like the topological consistency the completeness is not a measure that fulfils the purpose of adding complementary data to a dataset, as with the object 'pole'. It can only be used as an identification on the completeness of the contributions.

Thematic accuracy

The thematic accuracy is influenced by two factors that are important to mention. First, the user needs in which the children are subjected to objects defined by the Cadastre. This means that currently the user need of children towards the objects has not been taken into account e.g. the low rates of measures for the object-type 'bus station' because children used it to point out bus stops.

Second, the results are influenced by the inevitability that multiple objects are within the same building. This meant that buffers of reference data overlap and contributions fell within both buffers. This was especially the case with secondary schools and primary schools which were often placed adjacent to each other (e.g. the van Hasselt school). This biases the data because all the data could be 'correct'.

Quality indicators

The quality indictors used within the case- and pilot-study are purpose and usage. The votes are analysed within the case-study. The purpose of the project is to identify techniques and methods to assess the quality of VGI collected by schoolchildren in order to determine the objects 'correct' and 'incorrect'. This is successfully done by the matched- not matched equation as described in the following section.

The usage can be discussed in two parts, (i) how well does the data collection fits the user needs and (ii) can the collected data be used to assess the data on objects determined 'correct' or 'incorrect'.

A questionnaire was held at both the case-study as the pilot-study. Within the case-study also a central discussion was held with the schoolchildren after the lesson. This gave an indication from the point of view from the children about the lesson and the web viewer. Some interesting observations:

- Even though there are always some children who are less interested, most of the children gave enthusiastic feedback.
- From time to time the Web viewer was not responding if a contribution was added. Unfortunately, the cause for this problem was not found during the project, but it caused some frustration. The children had to wait until the web viewer was working again. Eventually this

influences the time that children had to make their contributions (yet another reason why the completeness measure was not suitable for the data assessment).

- The children could choose between 10 different POI. An extension of the number of objects or a bigger granularity could be interesting for the children e.g. bus station and bus stops. Another example of a useful extension would be to give the possibility is to point out that something does not exist anymore e.g. the situation displayed in figure 13. Contributions stated that there was a rail crossing. Because this was an identification to check the reference dataset, a 360° photograph was used. It turned out that the rail track was out of function.

Figure 13. Closed rail crossing

- There was no immediate feedback on the contributions, it was noticed that children would ‘have liked this, as the questioned was raised if we could tell them time how ‘good’ they had contributed.

Second, can the collected data be used for its purpose? The new equation matched- not-matched was developed specifically for this project. The data determined ‘correct’ is highly influenced by the distance calculated for the buffers. This makes the determination between ‘correct’ and ‘incorrect’ a dynamic parameter, which cannot be discussed without the buffers being known. The data collection within the case-study and pilot-study are assessed by this equation. Even though there was no reference data available for the object ‘pole’ the object has been compared with 360° photograph in which 6 of 18 were correct.

The equation is considered a useful measure to determine the ratio matched- and not-matched data, and so fulfilling the purpose needs on identifying ‘correct’ and ‘Incorrect data. It depends on the Cadastre how strict they define rules to identify what is ‘correct’ or ‘incorrect’. if the motivation is to compliment instead of complete or substitute a dataset, the data could be used as an indicator e.g. ‘poles’

Additional quality indicators

The additional quality indicators are the 'multiple contributions' and 'votes'. The 'multiple contributions' quality indicator is used within the case-study and the pilot-study. It turned out to be a useful indication of the objects missing in the reference dataset. If multiple contributions pointed out the same object, the reference dataset was checked and if necessary the missing object was added. Especially because the children could not see the contributions of others, this certainty that an object was missing increased.

This 'votes' quality indicator is only used within the case-study. We may assume that the high number of votes, even though the children only had 5 minutes, is due to the reason that voting costs less time. Additionally, it was noticed that children liked that the counter within the web viewer quickly went up.

In general, the 'positive' voted contributions have higher quality rates than the 'negative' voted contributions. However, what is interesting is that the 'negative' as well as the 'positive' voted contributions have high thematic accuracy rates.

While evaluating the contributions with the highest negative values so –4, -5 and -6 the likelihood that a contribution is incorrect increases. However, it is not determined to be incorrect. Within figure 14. the selected point has an averaged voting value of -6. Yet, based on the data-matching method the point is 'correct'. This means that it cannot be stated that from a certain amount of negative votes the contribution is likely to be incorrect. Nonetheless, it could give the indication that the contribution is likely to have a low precision.

Figure 14. Point selected vote -6

CHAPTER 6

Conclusion

The central research question of this master thesis is: “How can we assess and ensure the quality of VGI contributed by school children?” To answer this question, it was operationalized by splitting it op into four sub-question, which are answered below:

Question 1: Which quality measures are used regarding the assessment of VGI?

The methods to address this question, are a literature study, interviews and the assessment within the case-study.

Quality measures often used are derived from the ISO: 19157:2013 standard. However, since this quality standard is written for the quality assessment of well-structured, authoritative datasets, it does not suit the heterogeneous character of VGI very well. This is specifically the case for the quantitative (external) measures: completeness, topological consistency, and positional-, temporal-, and thematic-accuracy. In context to the case-study, these measures are calculated by measuring the discrepancy between the VGI dataset and the reference dataset. Some difficulties arise, such as the thematic differences between both datasets, the absence of reference datasets for specific object-types, and the differences in feature-types e.g. the VGI dataset only contained point-features referring to real-world entities, while the reference data represents these entities as point-, polygon-, and line-features. Additionally, the use of an authoritative data set to measure the quality of VGI runs contrary to the opportunity and motivation of using VGI to complement a dataset. Additionally, based on the believe that VGI can complement a dataset, the results of a comparison with the reference dataset could be biased because of the ‘omission’ or ‘commission’ in the reference dataset itself.

Contrasting to the quality measures, the quality indicators (purpose, usage, and lineage) provided the possibility to identify the internal quality of data without difficulties. In context of the case-study, the quality indicators seem to give reliable results since they are not dependent on external datasets.

Besides the ISO: 19157:2013 standard, two additional quality indicators were identified based on the interviews conducted in this research: (i) the ‘multiple contributions’ indicator to identify missing features in a reference dataset (ii) the ‘voting’ indicator to indicate the quality of the data. These quality indicators are successfully used within the quality assessment method.

Additionally, because the purpose of the project was to measure the rate 'correct' and 'incorrect' data, the new 'matched- not matched' equation was developed. Even though this measures was also calculated by the use of reference data, which gave the same difficulties as the external quality measures of the ISO standard, the equation provided the results expected to answer the purpose of the project.

Question 2: What are the possibilities and issues of VGI projects in the context of the Dutch national mapping agency?

The issues that holds NMAs back from using VGI are: the heterogeneous crowd and its unknown motivation, the unknown sustainability of projects, privacy- ethical-, and legal issues of using crowdsourced data, jobs replacement fears, and last but not least the quality of VGI.

Within this research, the quality of the data of children has been assessed mainly based on the 'matched- not matched' equation. Within the case-study 70% of the contributions could be matched and are therefore determined correct. The pilot-study used four different data collection methods in which Wageningen had a match of 100%, Apeldoorn 51.1%, Wolfheze 58.2% and Utrecht 68.4%. The calculated results depend on the data collection and data-matching method used within this research. In general, it depends on the opinion of the target user if the methods and results are both sufficient enough to decide of to make use of the data.

The Cadastre currently has no guidelines about the conditions that a VGI dataset must meet. It is likely that this will depend on the 'role' that the VGI will have e.g. if the purpose is to substitute or create a new dataset, the quality standard will probably have high restriction conform the quality standard used for authoritative sets. However, if the purpose is to complement or update a dataset it is believed that the collection of data by children could give possibilities in order to do so. Yet, since this research was a first step to perform a quality assessment, the possibilities and issues must be further explored.

Question 3 What are reliable quality assessment methods for VGI provided by school children?

Within this research, two quality methods are used. (i) the method developed to define the conditions of the case-study and the quality elements necessary to assess its quality (section 3.1), and (ii) an adjusted quality assessment method to assess the quality of the datasets collected within the pilot-study (section 3.2).

The first quality assessment method is developed to assess the VGI collected within the case-study. The method makes use of:

- The ISO quality measures: topological consistency (integrated in the web viewer), thematic accuracy, and completeness;

- The ISO quality indicators: purpose and usage in which the ‘matched not-matched’ equation was used;
- The additional quality indicators ‘multiple contributions’ and ‘votes’.

Based on the experience within the case-study, the general conclusion is that the reliability of this method depends on the intended purpose of the project. The purpose of the project was to identify the rate correct and incorrect features. The ISO quality measures could not assist in providing information for this purpose. For this reason, the matched not-matched equation was developed. This measure was successfully executed and resulted the information necessary to provide the rate.

The quality indicator ‘multiple contributions’ was used to ensure the quality of the reference dataset. Because the contributions were made separately without seeing the contributions of others, the reliability of the contributions increased by the increased number of objects referring to the same object. If these objects were not existing in the reference data they were check with aerial images and 360° photographs, and if correct added to the reference dataset.

The quality indicator ‘votes’ is not serving the immediate purpose of the project. However, it did provide interesting first results. As the relatively high number of votes and the potential of indicating the quality of data.

The method developed within the case-study is used to develop the method to assess the quality of the pilot-study. Some adjustments had to be made, because of differences within the data collection method. Additionally, the completeness measure has not been calculated, because the results of this measure was not suitable for the quality assessment. The measures are used within the same way as within the case-study. The following quality elements are used in the method:

- The ISO quality measures: thematic accuracy
- The ISO quality indicators: usage, and purpose (‘matched not-matched’);
- The additional quality indicators ‘multiple contributions’

The reliability of the results derived from both methods can still be questioned since the results are largely depending on the crowd contributing, the data-matching method used and the choice of POIs integrated within the web viewer. However, as foresaid to determine what the quality of the data ‘should’ be the role of the VGI must be clear. If the role is to complement or update a dataset the method is believed to be sufficient.

Question 4 What are requirements and strategies to integrate VGI contributed by children into NMA workflows?

The recommendation is based on the case-study and pilot-study assessed during this research. It is structured by the issues defined in the theoretical background.

nature of crowd and motivation

the requirement for the NMA is to have an idea of who is contributing. This can help to better contextualize the VGI. The 'Crowdsourcing at school!' is actually a good example of how this can be accomplished. During the visits an indication can be obtained from who contributed and why, and e.g. what are their recommendations to make the web viewer more interesting or user-friendly. An indication of their motivation to contribute, strongly relates to the sustainability of the project.

sustainability

Main opportunities of VGI within NMAs are to update existing datasets. A requirement is to have the project as a sustainable, on-going activity. It would be a waste of resources if it remains a one-time activity from which is presumed no NMA wants. A strategy to make the project sustainable is to learn more about the motivations of the contributors, and tailor the whole design of the VGI integration so that these motivations are addressed.

Privacy ethical and legal regulations

Because a VGI project is always a public data collection, privacy, ethical and legal regulations are inescapable to discuss. Concerning the privacy issue, the requirement is to gain more knowledge about the privacy regulations applicable within a VGI project. Currently the guiding principle is to collect as little private information as possible. However, if the project will become an on-going activity, it is advised to further research the privacy regulations. From the ethical perspective, it is recommended to have an overview of the ethical issues concerning the contributor's side as well as the project initiator side. The same can be said about the legal regulations concerning the property and liability. A strategy to address these issues is to do more research so that possibly a general mechanism can be developed.

Work issue

Main opportunities of VGI within NMAs are to update existing datasets, but it still needs to be validated by an NMA. A strategy to take away the fear of losing jobs is to share this knowledge e.g. organizing meetings, workshops and so on in which question and answers can be discussed.

Data quality

Because of the benefits which VGI could offer, the requirement is to do more research on the quality of VGI. The 'Crowdsourcing at school!' project can be further developed in order to do so. Within the recommendations to the research field in the 6.2. some ideas for further research are addressed.

Answer main research question: "How can we assess and ensure the quality of VGI contributed by school children?"

The quality assessment methods developed and used within this research are a first step towards a quality assessment method. Based on the results, more research is necessary to discover generic workflows or best practices that could be used to assess the VGI collected by schoolchildren.

6.1 Research limitations

Looking back at the research process different limitations have to be discussed:

- It was challenging to find enough schools to contribute. This meant that from the 34 schools in Hilversum in total. there were 3 positive responses. The remaining schools did not want to contribute or did not respond. The general answer was that the children have a busy schedule and that schools are often asked to contribute to projects in which they have to make choices. Because of this, the data collection is only done at these 3 schools.
- The results of the quality measure are based on small databases. We may assume that the collected data gives an indication on the quality of data collected by children. Still more school visits would enhance the knowledge on this assumption. The size of a desired sample group has not (yet) been discussed within scientific literature.
- Additionally, the available time for the project was limited. For this reason, the process of the research was accelerated. As for example to carry out visits at schools, the project was dependent on external factors e.g. the participation of schools and the dependency of school schedules. For this reason, appointments for visits had to be made early within the process. But before this could be done, the Web viewer had to be ready. This had put quite some time pressure on the development and adjustments necessary within the web viewer to conduct the case-study.

6.2 Recommendations research field

This research is a first step towards a quality assessment of VGI collected by schoolchildren. This leaves room for interesting follow-up studies:

- The quality measure thematic accuracy and completeness cannot directly be used to measure the correctness of the dataset. However, still they can assist within a quality assessment. The completeness measure can give an indication of the local knowledge of the children. For example, what is the number of objects found within the neighborhood or what is the range of the contributions and does this influence the correctness of objects. Within the research the children were subjected to time. However, in order to measure completeness, the children should be able to contribute as long as they would like to. Thematic accuracy can give an indication on the user needs to point out objects, as the object-type bus station which was often used to point out a bus stop.

- Additionally, to the thematic accuracy, it could be interesting to research the objects types that can be collected by children. For example, to relate more to their objects of interest e.g. playgrounds, football field, school etc. assuming that the children will be more interested and could have more knowledge about these kinds of objects.
- Another interesting idea for a follow-up study could be that children do not see an empty map, but that the map shows one or two object-types followed by a question on which object are correct and which ones are not., and if there are any objects missing. One could add a couple of false ones and hide a couple of known ones to make sure that there is something to detect. Award points could be given for detection of errors of commission and omission. Multiple of the same contributions could functions as triggers to analyse the data of the base data.
- The 'multiple contributions' indicator can be further asessed to identify rules that can be used to determine the reliability of data based on the number of contributions for the same object.
- The same can be said about the quality indicator votes which can give an indication on the quality of the data. Possibly these rules could be integrated into the web viewer e.g. if a contribution receives 10 negative votes within a X time the contributions are deleted automatically.

CHAPTER 7

Literature

Antoniou, V., Morley, J. and Haklay, M. (2009) 'The role of user generated spatial content in mapping agencies', *Proceedings of the {GIS} Research {UK} 19th …*.

Antoniou, V. (2011). User Generated Spatial Content: An Analysis of the Phenomenon and its Challenges for Mapping Agencies. *Doctoral Thesis, UCL (University College London).*, 324. Retrieved from http://discovery.ucl.ac.uk/1318053/

Antoniou, V., & Skopeliti, A. (2015). Measures and Indicators of Vgi Quality: an Overview. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II-3/W5*, 345–351. http://doi.org/10.5194/isprsannals-II-3-W5-345-2015

Antoniou, V., Christophe, S., Touya, G. and Skopeliti, A. (2016) 'Automated Mapmaking with Volunteered Geographic Information', *Mapping and the Citizen Sensor*, Chapter 4.

Bartoschek, T. and Keßler, C. (2012) 'VGI in education: From K-12 to graduate studies', in *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, pp. 341–360. doi: 10.1007/978-94-007-4587-2_19.

Coleman, D. J., Georgiadou, Y., Labonte, J., Observation, E., & Canada, N. R. (2009). Volunteered Geographic Information : The Nature and Motivation of Produsers. *International Journal of Spatial Data Infrastructures Research*, *4*(4), 332–358. http://doi.org/10.2902/1725-0463.2009.04.art16

Craglia, M., Ostermann, F., & Spinsanti, L. (2012). Digital Earth from vision to practice: making sense of citizen-generated content. *International Journal of Digital Earth*, *5*(5), 398–416. http://doi.org/10.1080/17538947.2012.712273

Devillers, R., & Roche, S. (2013). Assessing volunteered geographic information (VGI) quality based on contributors' mapping behaviours. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-2/W1*(June), 149–154. http://doi.org/10.5194/isprsarchives-XL-2-W1-149-2013

Dueren, M., & Bartoschek, T. (2013). Assessing the Usability of WebGIS for Schools. In *GI_FORUM 2013: CREATING THE GISOCIETY* (pp. 388–398). http://doi.org/10.1553/giscience2013s388

Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching Volunteered Geographic Information : Spatial Data , Geographic Research , and New Social Practice. *Annals of the Association of American Geographers*, *102*(3), 571–590. http://doi.org/10.1080/00045608.2011.595657

Euro SDR (2017). *Crowdsourcing in National Mapping 2017 - An International Workshop*. [information] Available at: http://www.cs.nuim.ie/~pmooney/eurosdr2017/

Estima, J., Pődör, A., Jokar, J., Laso-Bayas, J., & Vatseva, R. (2016). Mapping and the Citizen Sensor, *Sources of VGI for Mapping*, Chapter 2.

Fast, V. and Rinner, C. (2014) 'A Systems Perspective on Volunteered Geographic Information', *ISPRS International Journal of Geo-Information*, 3, pp. 1278–1292. doi: 10.3390/ijgi3041278.

Fritz, S., Fonte, C. C., Foody, G., Bastin, L., Olteanu-Raimond, A.-M., Mooney, P., Vatseva, R. (2016). *Mapping and the Citizen Sensor*, Chapter 1.

Fritz, S., See, L., Brovelli, M. and Campus, C. (2016) 'Motivating and Sustaining Participation in VGI', *Mapping and the Citizen Sensor*, chapter 5: sustaining and motivating in VGI

Fonte, C., Antoniou, V., Bastin, L., Laso-Bayas, J., See, L., & Vatseva, R. (2016*). Mapping and the citizen sensor*, Chapter 7: Assessing VGI Data Quality.

Forghani, M. and Delavar, M. (2014) 'A Quality Study of the OpenStreetMap Dataset for Tehran', *ISPRS International Journal of Geo-Information*, 3(2), pp. 750–763. doi: 10.3390/ijgi3020750.

GCN. (2017). *Worried about security? Beware the mosaic effect -- GCN*. [online] Available at: https://gcn.com/articles/2014/05/14/fose-mosaic-effect.aspx [Accessed 6 Feb. 2017].

Girres, J. F. and Touya, G. (2010) 'Quality Assessment of the French OpenStreetMap Dataset', *Transactions in GIS*, 14(4), pp. 435–459. doi: 10.1111/j.1467-9671.2010.01203.x.

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*. http://doi.org/10.1007/s10708-007-9111-y

Goodchild, M. F., & Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, *3*(3), 231–241 http://doi.org/10.1080/17538941003759255

Goodchild, M. F. and Li, L. (2012) 'Assuring the quality of volunteered geographic information', *Spatial Statistics*, 1, pp. 110–120. doi: 10.1016/j.spasta.2012.03.002.

Granell, C. and Ostermann, F. O. (2015) 'Beyond data collection: Objectives and methods of research using VGI and geo-social media for disaster management', *Computers, Environment and Urban Systems*. doi: 10.1016/j.compenvurbsys.2016.01.006.

Hashemi, P. and Abbaspour, R. A. (2015) 'Assessment of Logical Consistency in OpenStreetMap Based on the Spatial Similarity Concept', *OpenStreetMap in GIScience: Experiences, Research, Applications*, (JANUARY), pp. 1–20. doi: 10.1007/978-3-319-14280-7.

ISO. (2013). *ISO 19157:2013 - Geographic information -- Data quality*. [online] Available at: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32575 [Accessed 26 Oct. 2016].

INSIPRE, (2013). ANNEX III, section 3.5.5 [online] Available at: http://inspire-regadmin.jrc.ec.europa.eu/dataspecification/themes/oi/Chapter7.pdf [Accessed 7 Dec. 2016].

Jackson, S., Mullen, W., Agouris, P., Crooks, A., Croitoru, A., & Stefanidis, A. (2013). Assessing Completeness and Spatial Error of Features in Volunteered Geographic Information. *ISPRS International Journal of Geo-Information*, *2*(2), 507–530. http://doi.org/10.3390/ijgi2020507

Jonietz, D. and Zipf, A. (2016) 'Defining Fitness-for-Use for Crowdsourced Points of Interest (POI)', *ISPRS International Journal of Geo-Information*, 5(9), p. 149. doi: 10.3390/ijgi5090149.

Keßler, C., Trame, J. and Kauppinen, T. (2011) 'Tracking editing processes in volunteered geographic information: The case of OpenStreetMap', *Conference on Spatial Information Theory: …*, pp. 6–8. Available at: http://www.carsten.io/cosit11poster.pdf\nhttp://www.carsten.io/iope2011.pdf.

Keßler, C., Theodore, R. and Groot de, A. (2013) ‘Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap’, *Lecture Notes in Geoinformation and Cartography*, 2013–Janua, pp. 225–245. doi: 10.1007/978-3-319-00615-4.

Mooney, P., Rehrl, K. and Hochmair, H. (2013) ‘Action and interaction in volunteered geographic information: a workshop review’, *Journal of Location Based Services*, 7(4), pp. 291–311. doi: 10.1080/17489725.2013.859310.

Mooney, P., Juul, N., Alvanides, S., Kerle, N. and Observation, E. (2016) ‘Considerations of Privacy , Ethics and Legal Issues in Volunteered Geographic Information’, Chapter 6.

Mooney, P., Minghini, M., Laakso, M., Antoniou, V., Olteanu-Raimond, A.-M. and Skopeliti, A. (2016) ‘Towards a Protocol for the Collection of VGI Vector Data’, *ISPRS International Journal of Geo-Information*, 5(11), p. 217. doi: 10.3390/ijgi5110217.

Nam, T. and Pardo, T. A. (2011) ‘Conceptualizing smart city with dimensions of technology, people, and institutions’, in *Proceedings of the 12th Annual International Digital Government Research Conference on Digital Government Innovation in Challenging Times - dg.o ’11*, p. 282. doi: 10.1145/2037556.2037602.

Oksanen, J., Bergman, C., Sainio, J. and Westerholm, J. (2015) ‘Methods for deriving and calibrating privacy-preserving heat maps from mobile sports tracking application data’, *Journal of Transport Geography*. Elsevier B.V., 48, pp. 135–144. doi: 10.1016/j.jtrangeo.2015.09.001.

Ostermann, F. O., & Spinsanti, L. (2011). A Conceptual Workflow For Automatically Assessing The Quality Of Volunteered Geographic Information For Crisis Management. *Agile 2011*, 1–6.

Olteanu-Raimond, A., Antoniou, V., Fonte, C. C., Fonsesca, A., Grus, M., Harding, J., … Skopeliti, A. (2016). Mapping and the Citizen Sensor, Chapter 13: *VGI in National Mapping Agencies: Experiences and recommendations*.

Olteanu-Raimond, A. M., Hart, G., Foody, G. M., Touya, G., Kellenberger, T. and Demetriou, D. (2016b) ‘The Scale of VGI in Map Production: A Perspective on European National Mapping Agencies’, *Transactions in GIS*, 21(1), pp. 74–90. doi: 10.1111/tgis.12189.

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., … Rutzinger, M. (2016). Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced

Geographic Information. *ISPRS International Journal of Geo-Information*, *5*(5), 55. http://doi.org/10.3390/ijgi5050055

Turner, A. J. (2006) *Introduction to Neogeography*, *OReilly Short Cuts*. Available at: http://books.google.com/books?hl=en&lr=&id=oHgDv4feV-8C&oi=fnd&pg=PA24&dq=Introduction+to+Neogeography&ots=wYr7RESSeW&sig=HTani7IqwF0NQmiqf8jZigkFguQ.

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (Muki). (2016a). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, *8816*(June), 1–29. http://doi.org/10.1080/13658816.2016.1189556

V7f.eu. (2016). *School Crowd Sourcing*. [online] Available at: https://www.v7f.eu/school-crowd/school- crowd.html [Accessed 9 Nov. 2016].

Zielstra, D. and Hochmair, H. H. (2013) 'Positional accuracy analysis of Flickr and Panoramio images for selected world regions', *Journal of Spatial Science*, 58(2), pp. 251–273. doi: 10.1080/14498596.2013.801331.Appendices

Appendix I: reference features

reference objects ((in)correct, thematic accuracy, completeness)

Objects	Reference data	Additional source
Hospital	Thematic Layer: (VERBLIJFSOBJECT) Accommodation object: healthcare function Date:15-12-2015, 13-01-2016, 10-05-2016 Coordinate system: EPSG:28992, RD_New Source: BAG	Google
Gas station	Thematic Layer: (VERBLIJFSOBJECT) Accommodation object: industrial function Date:15-12-2015, 13-01-2016, 10-05-2016 Coordinate system: EPSG:28992, RD_New Source: BAG	https://www.anwb.nl/auto/brandstof/tankstations and 360° photographs
Primary schools	Thematic layer: (BASIS SCHOLEN) Date: 15-05-2016 Coordinate system: EPSG:28992, RD_New Source: Datasets- ESRI Nederland	www.scholenopdekaart.nl and 360° photographs*
High schools	Thematic layer: (MIDDELBARE SCHOLEN) Date: 15-05-2016 Coordinate system: EPSG:28992, RD_New Source: Datasets- ESRI Nederland	www.scholenopdekaart.nl and 360° photographs *
Police stations	Thematic layer: (POLITIE BUREAUS) Date: 06-05-2013 Coordinate system: EPSG:28992, RD_New Source: Datasets- ESRI Nederland	Google and 360° photographs*
Fire stations	Thematic layer (BRANDWEERKAZERNES) Date: 06-05-2013 Coordinate system: EPSG:28992, RD_New Source: Datasets- ESRI Nederland	Google and 360° photographs*
Rail crossing	Thematic Layer Date: 01-12-2015, 27-02-2016 Coordinate system: EPSG:28992, RD_New Source: TOP10NL	PDOK(WMS)
Bus stations	Thematic Layer Date: 28-01-2016 Coordinate system: EPSG:28992, RD_New Source: Geo service university Utrecht Area: The Netherlands	Aerial imagery
Train Stations	Thematic layer(STATIONS) Date: 07-07-2006 Coordinate system: EPSG:28992, RD_New Source: Geo service university Utrecht Area: The Netherlands	Aerial imagery
Pole	-	360° photographs

*The 360° photographs are derived from Globe spotter, Cyclomedia, dated: 22-12-2016

* the aerial imagery are derived from PDOK(2015)

Appendix II summary interviews

Questions and summarized answers of the interviews with Martin Kraan, Jasper Hogerwerf, Ana-Maria Olteanu-Raimond, and Magdalena Grus.

Interview: Martin Kraan (face-to-face)

Can you explain the VGI project which you are working on?

Martin Kraan is working on a successful VGI project released on 23 November 2016 called the ‘feedback system’. This is a ‘change detection’ and ‘error reporting’ type of VGI use.

The crowd can add their contributions within a web viewer in which the BRT is displayed. While placing the contribution, there is an immediate feedback that the contribution is pending to be checked which will normally takes 1-2 days. If the checking is done, the contributor receives feedback about what has been done with his or her add.

How many contributions do you receive with the new system?

Before the feedback system, the step to contribute was difficult. The contributor needed log in codes and they were obligated to fill in an email address. This resulted in 1-2 contributions per month. With the feedback system, these administrative rules were no longer necessary e.g. they can still fill in their email address but this is not obligatory. Currently they receive 20-30 contributions on average per month (at the beginning it was more than 300).

How do you assess if a contribution is either correct or incorrect?

These checking’s are done by a few people working at the Cadastre, Maarten Kraan is one of these people. To determine if a contribution is either correct or incorrect they compare the contributions with recent aerial imagery or if necessary 360° photographs. If the contributor is correct an adjustment is made in the dataset otherwise it is not. The adjustments will be present in the following release.

Do you assess the quality of the contributions? For instance, with the ISO standards?

No quality assessment is done based on the contributions. The only assessment within this system is the comparison explained in the previous question.

Do you measure the reliable are the contributions? And how reliable are the contributors?

There is no system to measure the reliability. Most of the contributions are placed correct. If they are incorrect it looks like the contributor did some trial and error within the web viewer.

Within my research I gather data by the use of crowdsourcing done by children. How do you think the data gathering should be visualized (integrated) within a web viewer?

The datasets can be integrated into the web viewer e.g. a school layer within the web viewer. If the contribution is correct the children can receive an immediate feedback on their contributions.

How do you think the quality of the contributions van be asessed?

With comparing the contributions with aerial images and 360 photographs.

Are there guidelines describing a minimum quality standard that crowdsourced data needs to have? i.e. Could VGI possibly be integrated in the SDI of the Cadastre when meeting certain conditions?

Martin thinks there is a no difference between the contributions of adults and the contributions of children. The contributions could possibly be integrated into the 'feedback system'.

Concluding

The 'feedback system' helps to increase the quality of the geo datasets. However, it does not give information on an assessment of the quality or guidelines of the contributions itself.

Interview: Jasper Hogerwerf (face- to-face)

Can you explain how VGI project which you are working on?

Jasper Hogerwerf is working on the toponyms pilot-study. 360° photographs can be used to check if the names of e.g. buildings and streets are still correct. However, because the names are often small it is difficult or impossible to read them. They asked different municipalities and historic associations to complement or update their datasets on toponyms.

How many contributions do you receive with the new system?

This pilot-study is still going. Half October first request were send and first results are received. There were two forms send back from the historic association that seemed 'good' in terms of completeness of information given.

How do you assess if a contribution is either correct or incorrect?

The contributions still need to be validated by a topographer. This is currently done by reviewing the possibilities of external sources such as the 360° photographs and information found on the internet. He or she will give feedback and if necessary make the adjustment within the datasets if the contributions are determined 'correct'.

Do you assess the quality of the contributions? For instance, with the ISO standards?

No not within this stage of the project. Currently is the assessment done as described within the previous question. Besides it is a question if a quality assessment based on these measures is applicable to the project.

Do you measure the reliable are the contributions? And how reliable are the contributors?

Idem.

Within my research I gather data by the use of crowdsourcing done by children. How do you think the data gathering should visualized(integrated) within a web viewer?

A method could be to first compare aerial imagery to filter out mutations which can functions as triggers. A following question could be what kind of objects can be collected by children. The contributions itself can be validated by aerial imagery, 360° pictures and the internet as an external source. A website however cannot always be trusted; this should be kept in mind.

Additionally, the question is if the objects(POIs) are interesting or applicable for children e.g. in Limburg there are a lot of crosses adjacent to the road (http://veldkruus.nl/). But there should be no situations in which children go walking on or adjacent to the road to find these objects. This can cause dangerous situations. On the other hand, it could be interesting to review per municipality were the need is for certain object e.g. Limburg has different objects and needs then Friesland.

How do you think the quality of the contributions van be asessed?

For the 'crowdsourcing at school project' The ISO standards can be analysed to determine which measure are applicable to the project as well as the data collection method within the web viewer which can function as a validation. A validation could be that if within the Web viewer children cannot see the contributions of other children, but multiple children say that an object is on the same place they are in a way validating each other.

Are there guidelines describing a minimum quality standard that crowdsourced data needs to have? i.e. Could VGI possibly be integrated in the SDI of the Cadastre when meeting certain conditions?

Not known

Concluding

The Toponyms project is aiming to increase the quality of the toponyms datasets rather than to analyse the quality of the contributions itself. Within the 'crowdsourcing at school project' the ISO quality measures can assist to determine the quality of data. However, the integration of validation methods into the web viewer would be interesting idea.

Interview Ana-Maria Olteanu-Raimond (skype and email contact)

The combination of measuring quality of VGI within an NMA is something I find difficult. Can you tell me more about how the France NMA coops with the quality assessment of VGI?

The idea is to compare homologous features (i.e. features that represent the same object in the real world). Thus, the first step is to identify homologous features. The process that allows defining homologous features is named data-matching. The data-matching can be done manually if your dataset is not so big, using buffers or using different methods of data-matching that are developed in literature.

positional accuracy, Thematic accuracy and completeness only makes sense if reference data is available.

- For measuring the accuracy, the process that I used is as follows: Once you identify the homologous features, you need to compare them in order to compute the accuracy. For doing that, we usually use metric measures: Euclidian distance for points, Hausdorff or Frechet distance for lines, and Surface distance (the ratio between the overlapped area and the union of two polygons; a measure between 0 and 1; 0 means that two polygons do not intersect; 1 means that they are completely overlapped) or angular distance for polygons.
- Completeness only makes sense if it has a meaning so if this is for example one of the goals obstacle. For objects as the 'pole' object it is not possible to give useful information.
- The thematic accuracy depends on the granularity of objects that is applied.
- Concerning the logical consistency, we developed different indicators that satisfy rules: for example, a road should not cross a building; a house should not be in a water body. It really depends on your dataset. These measures are easy to implement in a GIS system.

I find it difficult to determine the buffers width per objects e.g. school buildings, gas station. This is the same within logical consistency I find it difficult to determine what is 'correct' or 'incorrect'. Do you know how this can be determined or if there are international rules written about this in context to the quality standards of NMAs?

There are no international rules; it really depends on your data. Additionally, it is important to determine which roll you have e.g. if the data is used as a trigger then it is different from data that will be integrated into the datasets.

Concluding

Ana-Maria provided interesting information about how VGI datasets can be compared with reference datasets during data-matching processes. Quality measures are project specific e.g. positional accuracy, Thematic accuracy and completeness makes sense if reference data is available, Topological rules can be added number as the number of violations of the topological scheme.

Interview: Magdalena Grus (face-to-face and email contact)

Nearly every week we had a meeting at the Cadastre to discuss decision made within this research. For this reason, no clear question and answers can be written down. However, concluding information relevant for the case-study is described.

Concluding

Magdalena has worked and is working on several projects at the Cadastre concerning the use of VGI, as the ‘crowdsourcing at school’ project. According to her the project was considered a success based on the response of the teachers and children. However, the quality assessment is according to her a big issue since there is no method yet at the Cadastre to assess the quality. According to her more research needs to be done on the quality assessment and the possibilities of introducing crowdsourcing at school. An interesting quality measure for her is the total number of features that is either ‘correct’ or ‘incorrect’.

References datasets that are available to compare the VGI maps with reference data are the: BRT, BAG, aerial imagery (10cm) and 360° photographs (Globe spotter). A method to match the data within a reliable way could be the following example: there is no authoritative data on the addresses of gas station. By consulting external sources, the addresses can be conducted. With aerial imagery, it is still difficult to define based on a roof (if there is a roof) if a place is a gas station. Within this case 360° photographs can be examined to see if there is a gas station.

An interesting quality indicator could be assessed by integrating a voting system into the Web viewer, it could serve as a reward, and it indicates how ‘good’ children can judge the contributions of others.

Appendix III: Web viewer

Screen 1) Fill in age

Screen 2) adding points of interest

Screen 3) see and vote

Screen 4) final screen and overview contribution

APPENDIX IV: Scheme of the teaching method

Time lead	What?	tasks	Who?	Where?	Tools
5 min.	**Introduction**	**Explanation of today**	**Teacher and Elisa**	**Classroom**	
5 min.	**What is crowdsourcing?**	**Introduction crowdsourcing with movie (2.35 min.)** https://www.youtube.com/watch?v=SGkzL6qWWOo&feature=youtu.be	**Elisa**	**Classroom**	**Digi board**
5 min.	**Short discussion after movie**	**Discuss questions** **Other examples of crowdsourcing?** **Why is crowdsourcing important for a map?**	**Elisa**	**Classroom**	
5 min.	**Explanation application**	**Explanation application** https://school-crowd.v7f.eu/intro.html	**Elisa**	**Classroom**	**Digi board**
15 min.	**Work!**	**-working with the web viewer (10 min. contributing/5 min. voting)** **-children in groups of 2-3 person**	**Elisa and Tom/ Stéphane/ Magdalena**	**Classroom/ computer room**	
5 min.	**Final discussion**	**- thank the children** **- ask for questions and possible improvements**	**Elisa**	**Classroom**	
In total: 40 min.					

Appendix V

Completeness school per group

School 1 Minckelerschool group 8

Points of interest	Total number of reference features	Number of objects found by the children	Completeness in percentages
Primary School	33	3	9.1
Secondary school	13	2	15.4
gas station	15	2	13.3
Police station	1	1	100
Hospital	1	1	100
Train stations	4	1	25
Rail crossing	5	3	60
Bus station	1	1	100
Fire station	1	1	100
Pole	-	-	-
Total	74	15	20.2

School 2 Van Hasselt school group 8

Points of interest	Total number of reference objects	Number of objects found by the children	Completeness in percentages
Primary School	33	3	9.1
Secondary school	13	2	15.4
gas station	15	1	6.7
Police station	1	0	0
Hospital	1	0	0
Train stations	4	0	0
Rail crossing	5	1	20
Bus station	1	1	100
Fire station	1	0	0
Pole	-	-	-
Total	74	8	10.8

School 3 Van Hasseltschool group 7

Points of interest	Total number of reference objects	Number of objects found by the children	Completeness in percentages
Primary School	33	1	3
Secondary school	13	2	15.4
gas station	15	1	6.7
Police station	1	1	100
Hospital	1	0	0
Train stations	4	0	0
Rail crossing	5	0	0
Bus station	1	0	0
Fire station	1	0	0
Pole	-	-	-
Total	74	5	6.8

School 4 Violenschool group 8a

Points of interest	Total number of reference objects	Number of objects found by the children	Completeness in percentages
Primary School	33	5	15.2
Secondary school	13	4	30.7
gas station	15	2	13.4
Police station	1	0	0
Hospital	1	0	0
Train stations	4	1	25
Rail crossing	5	3	60
Bus station	1	0	0
Fire station	1	1	100
Pole	-	-	-
Total	74	16	21.6

School 5 Violenschool 8b

Points of interest	Total number of reference objects	Number of objects found by the children	Completeness in percentages
Primary School	33	1	3
Secondary school	13	7	53.8
gas station	15	3	20
Police station	1	1	100
Hospital	1	1	100
Train stations	4	1	25
Rail crossing	5	2	40
Bus station	1	1	100
Fire station	1	0	0
Pole	-	-	-
Total	74	17	22.9

Results in total

Points of interest	Total number of reference objects	Number of objects found by the children	Completeness in percentages
Primary School	33	13	39.4
Secondary school	13	11	84.6
gas station	15	6	40
Police station	1	1	100
Hospital	1	1	100
Train stations	4	2	50
Rail crossing	5	1	20
Bus station	1	1	100
Fire station	1	1	100
Pole	-	-	-
Total	74	37	50

Appendix VI

Thematic accuracy

School 1 Minckelerschool group 8

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	10	10	100
Secondary school	4	4	100
Gas station	2	2	100
Police station	3	3	100
Hospital	4	4	100
Train stations	8	5	62.5
Rail crossing	3	3	100
Bus station	3	2	66.7
Fire station	1	1	100
Pole	-	-	-
Total	38	34	89.5

School 2 Van Hasselt school group 8

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	11	8	72.7
Secondary school	11	5	45.5
Gas station	6	6	100
Police station	0	0	0
Hospital	0	0	0
Train stations	1	0	0
Rail crossing	1	1	100
Bus station	1	1	100
Fire station	0	0	
Pole	-	-	-
Total	31	21	67.7

School 3 Van Hasseltschool group 7

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	14	8	57.1
Secondary school	15	6	40
Gas station	4	4	100
Police station	1	1	100
Hospital	0	0	0
Train stations	0	0	0
Rail crossing	0	0	0
Bus station	0	0	0
Fire station	0	0	0
Pole	-	-	-
Total	34	19	55.9

School 4 Violenschool group 8a

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	18	13	72.2
Secondary school	17	13	76.4
Gas station	4	4	100
Police station	1	1	100
Hospital	0	0	0
Train stations	5	3	60
Rail crossing	7	7	100
Bus station	1	0	0
Fire station	2	2	100
Pole	-	-	-
Total	55	43	78.2

School 5 Violenschool 8b

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	9	8	88.9
Secondary school	12	11	91.6
Gas station	5	5	100
Police station	1	1	100
Hospital	4	3	75
Train stations	7	2	28.6
Rail crossing	4	4	100
Bus station	2	2	100
Fire station	0	0	0
Pole	-	-	-
Total	44	36	81.8

Results in total

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	62	47	75.8
Secondary school	59	39	66.1
Gas station	21	21	100
Police station	6	6	100
Hospital	8	7	87.5
Train stations	21	10	47.6
Rail crossing	15	15	100
Bus station	7	5	71.4
Fire station	3	3	100
Pole	-	-	-
Total	202	153	75.7

Appendix VII

Matched proximity and thematically

School 1 Minckelerschool group 8

Points of interest	VGI total	Matched data	Percentage correct
Primary School	12	10	83.3
Secondary school	4	4	100
Gas station	2	2	100
Police station	3	3	100
Hospital	4	4	100
Train stations	6	5	83.3
Rail crossing	4	3	75
Bus station	2	2	100
Fire station	3	1	33
Pole	2	2	100
Total	42	36	85.7

School 2 Van Hasselt school group 8

Points of interest	VGI total	Matched data	Percentage correct
Primary School	10	8	80
Secondary school	10	5	50
Gas station	6	6	100
Police station	1	0	0
Hospital	0	0	0
Train stations	1	0	0
Rail crossing	1	1	100
Bus station	2	1	50
Fire station	0	0	0
Pole	0	0	0
Total	31	21	67.7

School 3 Van Hasseltschool group 7

Points of interest	VGI total	Matched data	Percentage correct
Primary School	9	8	88.9
Secondary school	7	6	85.7
Gas station	5	4	80
Police station	1	1	100
Hospital	2	0	0
Train stations	2	0	0
Rail crossing	2	0	0
Bus station	1	0	0
Fire station	0	0	0
Pole	3	0	0
Total	32	19	59.4

School 4 Violenschool group 8a

Points of interest	VGI total	Matched data	Percentage correct
Primary School	13	13	100
Secondary school	18	13	72.2
Gas station	4	4	100
Police station	2	1	50
Hospital	1	0	0
Train stations	4	3	75
Rail crossing	8	7	87.5
Bus station	2	0	0
Fire station	3	2	66.7
Pole	7	2	28.6
Total	62	45	72.6

School 5 Violenschool 8b

Points of interest	VGI total	Matched data	Percentage correct
Primary School	12	8	66.7
Secondary school	14	11	78.6
Gas station	5	5	100
Police station	4	1	25
Hospital	3	3	100
Train stations	3	2	66.7
Rail crossing	6	4	66.7
Bus station	6	2	33
Fire station	1	0	0
Pole	6	2	33.3
Total	60	38	63.3

Results in total

Points of interest	VGI total	Matched data	Percentage correct
Primary School	56	47	83.9
Secondary school	53	39	73.6
Gas station	22	21	95.5
Police station	11	6	54.5
Hospital	10	7	70
Train stations	16	10	62.5
Rail crossing	21	15	71.4
Bus station	13	5	38.5
Fire station	7	3	42.8
Pole	18	6	50
Total	227	159	70

Appendix VIII

Results in total votes (left positive and correct the negative valued contributions)

Match proximity and thematical (positive)

Points of interest	VGI total	Matched data	Percentage correct
Primary School	33	30	90.9
Secondary school	22	21	95.5
Gas station	18	17	94.4
Police station	2	0	0
Hospital	2	1	50
Train stations	7	6	85.7
Rail crossing	8	6	75
Bus station	4	2	50
Fire station	2	1	50
Pole	4	2	50
Total	102	86	84.3

Match proximity and thematical) (negative)

Points of interest	VGI total	Matched data	Percentage correct
Primary School	8	2	25
Secondary school	20	12	60
Gas station	0	0	0
Police station	6	3	50
Hospital	4	3	75
Train stations	7	3	42.8
Rail crossing	2	1	50
Bus station	4	0	0
Fire station	2	1	50
Pole	9	3	33.3
Total	62	28	45.1

Thematic accuracy (positive)

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	38	30	78.9
Secondary school	31	21	67.7
Gas station	17	17	100
Police station	0	0	0
Hospital	1	1	100
Train stations	10	6	60
Rail crossing	6	6	100
Bus station	3	2	66
Fire station	1	1	100
Pole	-	-	-
Total	107	84	78.5

Thematic accuracy (negative)

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	6	2	33.3
Secondary school	14	12	85.7
Gas station	0	0	0
Police station	3	3	100
Hospital	4	3	75
Train stations	4	3	75
Rail crossing	1	1	100
Bus station	1	0	0
Fire station	1	1	100
Pole	-	-	-
Total	34	25	73.5

Appendix IX

Thematic accuracy Pilot-study

Thematic accuracy (Wageningen)

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	2	2	100
Secondary school	0		
Gas station	0	-	-
Police station	1	1	100
Hospital	1	1	100
Train stations	0	-	-
Rail crossing	0	-	-
Bus station	0	-	-
Fire station	0	-	-
Pole	0	-	-
Total	4	4	100

Thematic accuracy (Apeldoorn)

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	16	14	87.5
Secondary school	5	5	100
Gas station	6	6	100
Police station	1	1	100
Hospital	2	2	100
Train stations	9	8	88.9
Rail crossing	7	7	100
Bus station	1	1	100
Fire station	1	1	100
Pole	-	-	-
Total	48	45	93.7

Thematic accuracy (Wolfheze)

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	9	9	100
Secondary school	1	1	100
Gas station	2	2	100
Police station	0	-	-
Hospital	0	-	-
Train stations	29	9	31
Rail crossing	29	12	41.3
Bus station	0	-	
Fire station	9	6	66.7
Pole	-	-	-
Total	79	39	49.4

Thematic accuracy (Utrecht)

Reference objects	VGI matched on proximity	VGI match thematically	Percentage
Primary School	16	9	56.3
Secondary school	22	22	100
Gas station	5	5	100
Police station	0	0	0
Hospital	9	8	88.9
Train stations	3	3	100
Rail crossing	2	2	100
Bus station	0	0	0
Fire station	0	-	-
Pole	-	-	-
Total	57	49	85.9

Appendix X

Matched- not-matched Pilot-study

Match proximity and thematical (Wageningen)

Points of interest	VGI total	Matched data	Percentage correct
Primary School	2	2	100
Secondary school	0	-	-
Gas station	0	-	-
Police station	1	1	100
Hospital	0	0	0
Train stations	0	-	-
Rail crossing	0	-	-
Bus station	0	-	-
Fire station	0	-	-
Pole	0	-	-
Total	3	3	100

Match proximity and thematical (Apeldoorn)

Points of interest	VGI total	Matched data	Percentage correct
Primary School	18	14	77.7
Secondary school	10	5	50
Gas station	8	6	75
Police station	6	1	16.6
Hospital	3	2	66.7
Train stations	11	8	72.7
Rail crossing	12	7	58.3
Bus station	7	1	14.3
Fire station	2	1	50
Pole	-	-	-
Total	77	45	51.1

Match proximity and thematical (Wolfheze)

Points of interest	VGI total	Matched data	Percentage correct
Primary School	11	9	81.8
Secondary school	4	1	25
Gas station	2	2	100
Police station	2	0	0
Hospital	2	0	0
Train stations	9	9	100
Rail crossing	12	12	100
Bus station	14	0	-
Fire station	8	6	75
Pole	-	-	-
Total	67	39	58.2

Match proximity and thematical (Utrecht)

Points of interest	VGI total	Matched data	Percentage correct
Primary School	14	9	64.3
Secondary school	23	22	95.6
Gas station	5	5	100
Police station	1	0	0
Hospital	9	8	88.8
Train stations	3	3	100
Rail crossing	2	2	100
Bus station	3	0	0
Fire station	0	-	-
Pole	16	3	18.8
Total	76	52	68.4