Quality Assurance in Geodata

Soran Sulaiman
Halldora Gudmundsdottir

Degree of Master Thesis (1y)
Stockholm, Sweden 2013
TMT 217
Abstract

The use of geodata is increasing all over the world and consequently data quality is receiving a higher priority. Nowadays, geodata organizations are putting more effort into analyzing their current methods of ensuring and maintaining data quality in order to meet the growing demands of customers. The Swedish government has placed a great emphasis on cooperation between organizations and launched a project for establishing a national infrastructure for geodata. For such a collaboration to be successful the reliability of produced geodata has to be high and accepted level of data quality to be ensured. The main objective of this study is to analyze the current data quality assurance processes of selected geodata organizations in Sweden (Lantmäteriet, Stockholms Stad and Sjöfartsverket), find disconnections and suggest improvements. Furthermore, a comparison is made with a data quality assurance process at an international organization, iMMAP.

The approach used for collecting data in this thesis was on-site interactive and qualitative interviews. During the interviews key personnel were present and provided an overview of the organization in question, its goals, operations and current QA processes and procedures.

The theoretical research performed in this study and the interviews emphasized the importance of data quality in the organizations. Another topic of high determination discovered during the interviews was the interest for improving their QA process, which is now integrated into the entire data management system. As a result, an effective quality assurance process is designed, mapped and recommended, through scrutinizing different methods of performing the process in the organizations. Moreover, development of a clear quality policy along with the organization main policy is advised.
Acknowledgements

We would like to thank Mr. Roland Langhé the director of Project Management and Operational Development Master Program at KTH for his support and guidance, especially at the initial stage of the thesis. Sincere gratitude is extended to Dr. Jesper Paasch, researcher and developer at Lantmäteriet and our supervisor, for his time, cooperation, patience, quick responses, wise guidance and helpful suggestions throughout this study.

Furthermore, we would like to express gratitude to all the personnel that took part in the interviews at the organizations, especially our contact persons;

- Dr. Jesper Paasch at Lantmäteriet
- Mr. Krister Hedman at Stockholms Stad
- Mr. Ralf Lindgren at Sjöfartsverket

Our high appreciation goes to iMMAP for their kind contribution to this thesis by providing us with a broader perspective on QA processes, through involvement of Mr. Bekim Kajtazi.

Certainly, acknowledgements are given to Geodata Secretariat (Geodatasekretariatet) at Lantmäteriet, which funded this research.
# Table of Contents

1  Introduction .........................................................................................................................1
   1.1  Background ......................................................................................................................1
   1.2  Present Situation .............................................................................................................1
   1.3  Goal ................................................................................................................................2
   1.4  Scope ................................................................................................................................2
   1.5  Methodology ....................................................................................................................2

2  Theoretical Reflection ..........................................................................................................7
   2.1  Geographic data and GIS .................................................................................................7
   2.2  Quality Control and Quality Assurance ..........................................................................10
   2.3  Planning for QA in Geodata ...........................................................................................10
   2.4  Data Quality and Uncertainties in Geographic Data .........................................................11
   2.5  Principles of Data Quality ...............................................................................................12
   2.6  Systematic QA Process Management and Process Mapping ........................................14

3  Fact Gathering ....................................................................................................................17

4  Description of Fieldwork .....................................................................................................19

5  Analysis of Data and Information .......................................................................................21
   5.1  Lantmäteriet ....................................................................................................................21
   5.2  Stockholms Stad ..............................................................................................................27
   5.3  Sjöfartsverket .................................................................................................................31
   5.4  iMMAP ..........................................................................................................................33
   5.5  Comparison .....................................................................................................................35

6  Recommendations ...............................................................................................................37

7  Conclusions ..........................................................................................................................41

References ....................................................................................................................................43

Appendix A ..............................................................................................................................45
Appendix B ..............................................................................................................................48
Appendix C ..............................................................................................................................50
Appendix D ..............................................................................................................................52
List of Figures

Figure 1. Typical Thematic Layer Model in GIS (GeoVITe, 2010) ........................................8
Figure 2. Important Components of Geodata: Space, Time and Topic (GeoVITe, 2010) .........8
Figure 3. Components of GIS (GeoVITe, 2010) ....................................................................9
Figure 4. A simple process map for monitoring failures at Lantmäteriet (Paasch, J., personal communication, 10th April 2013) ........................................................................24
Figure 5. The IS map for analyzing and managing quality failures in BAL at Lantmäteriet (Paasch, J., personal communication, 25th April 2013) .................................................................26
Figure 6. Flow Chart for Data Production Environment (Hedman, K., personal communication, 13th May 2013) ........................................................................................................30
Figure 7. Data Updating Process Chain (Hedman, K., personal communication, 13th May 2013) .................................................................................................................................30
Figure 8. General Data Flow in Sjöfartsverket (Lindgren, R., personal communication, 15th May 2013) ..................................................................................................................33
Figure 9. Process Map for iMMAP Information Management System (Kajtazi, B., personal communication, 9th May 2013) ..............................................................................................34
Figure 10. The Should Map for the data QA process .............................................................39
Figure 11. Continuous QA Process for Data Management Systems .....................................40
Figure 12. Main process map for information management at Lantmäteriet (Paasch, J., personal communication, 16th April 2013) ..............................................................54
Figure 13. The Process for receiving, verifying and updating data about ABT (Paasch, J., personal communication, 16th April 2013) .................................................................55
Figure 14. The support process for buildings, addresses and apartments (BAL) (Paasch, J., personal communication, 16th April 2013) .................................................................56

List of Tables

Table 1. Popular geodata usage groups and scenarios (GeoVITe, 2010) .................................9
Table 2. Quality and GEO information related ISO standards. Information retrieved from ISO’s (International Organization for Standardization) official website www.iso.org ..........14
Table 3. Essential factors about geodata quality (Yang, n.d.) ..................................................47
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABT</td>
<td>Address, Building, Topology (Adress, Byggnad, Topologi)</td>
</tr>
<tr>
<td>AIS</td>
<td>Automatic Identification System</td>
</tr>
<tr>
<td>BAL</td>
<td>Building, Address, Apartment (Byggnad, Adress, Lägenhet)</td>
</tr>
<tr>
<td>CL</td>
<td>Clearance</td>
</tr>
<tr>
<td>DP</td>
<td>Digital Photo</td>
</tr>
<tr>
<td>GDS</td>
<td>Base Data System (GrundDataSystem)</td>
</tr>
<tr>
<td>Geodata</td>
<td>Geographic data</td>
</tr>
<tr>
<td>GGD</td>
<td>Basic geographical data (Geografiska GrundData)</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographical Information System</td>
</tr>
<tr>
<td>GNSS</td>
<td>Global Navigation Satellite Systems</td>
</tr>
<tr>
<td>HMK</td>
<td>Handbook for Measurement and Mapping Matters (Handbok i mät och kartfrågor)</td>
</tr>
<tr>
<td>IFS</td>
<td>Information Management Support (Informationsförvaltningsstöd)</td>
</tr>
<tr>
<td>IHO</td>
<td>International Hydrographic Organization</td>
</tr>
<tr>
<td>IM</td>
<td>Information management</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>iMMAP</td>
<td>Information Management and Mine Action Programs</td>
</tr>
<tr>
<td>IMSMA</td>
<td>Information Management System for Mine Action</td>
</tr>
<tr>
<td>LevFR</td>
<td>Delivery Data Store (Tillhandahållandelan)</td>
</tr>
<tr>
<td>LINA</td>
<td>Lantmäteriet data collection application (Lantmäteriet InSamlingsApplikation)</td>
</tr>
<tr>
<td>LV</td>
<td>Local Road Network Database (Lokal vägnät databas)</td>
</tr>
<tr>
<td>MRE</td>
<td>Mine Risk Education</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-Governmental Organization</td>
</tr>
<tr>
<td>OPS</td>
<td>Operations</td>
</tr>
<tr>
<td>Sigma</td>
<td>Property Delivery Store (Fastighet Tillhandahållandelager)</td>
</tr>
<tr>
<td>SWEPOS</td>
<td>A network of fixed GNSS stations</td>
</tr>
<tr>
<td>TP</td>
<td>Turning point</td>
</tr>
<tr>
<td>QA</td>
<td>Quality assurance</td>
</tr>
<tr>
<td>QC</td>
<td>Quality control</td>
</tr>
<tr>
<td>QF</td>
<td>Quality failure</td>
</tr>
<tr>
<td>VA</td>
<td>Victim Assistance</td>
</tr>
<tr>
<td>VTS</td>
<td>Vessel Traffic Service</td>
</tr>
<tr>
<td>VVAF</td>
<td>Vietnam Veteran of America Foundation</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Background

Geographical information has become crucial for sound decision making in today’s society at local, sectional and global levels. It is an essential part of a nation’s digital information infrastructure and benefits decision makers in many areas such as in flood mitigation, disaster recovery, crime management, urban and rural planning, environmental reconstruction and route planning for ambulance, police and firefighting services. Although much is to be gained using geographical information in social planning and decisions making, this information is not always readily available, up-to-date or accurate enough. Furthermore, information is an expensive resource and great efforts are needed for collecting, analyzing and assuring quality. Many programs on all levels, local, regional, national and global, are working towards developing greater access to geographical information and ensuring its quality and up-to-date status. It is the general opinion that adjusting standards and coordinating data collection and maintenance for various agencies, organizations and companies will result in better work practices and produced data (SDI, 2012).

The subject of this study is to analyze the management of geographical information and how data is kept in accordance with specifications, regarding accuracy, actuality and completeness. The main objective is to investigate the quality assurance (QA) processes of selected national and municipal organizations, find disconnections in the processes if there are any and suggest improvements. More closely, the goal is to examine how organizations are guaranteeing and maintaining the quality of the information they produce, buy or use within their respective region. By analyzing different QA processes and comparing how organizations manage their data quality issues valuable knowledge may be gained. This knowledge can be used to integrate geographical data from various organizations which will in turn enhance the reliability and credibility of data produced.

In this study the focus is on analyzing the geodata QA process within Lantmäteriet, a Swedish mapping, cadastral and land registration authority, Stockholms Stad, the City of Stockholm, and Sjöfartsverket, the Swedish Maritime Administration. A comparison is made between the QA in these organizations in order to identify disconnections and improvement measures. Furthermore, the QA process in an international organization, iMMAP, is discussed briefly. This is done in order to gain a broader perspective at the issue at hand and to be able to suggest well-informed QA process improvements.

1.2 Present Situation

The value and quality of geographical information are ever increasing. Nowadays, the demand for more complex data, increased speed to market and frequent data updates is high and therefore means for ensuring quality are becoming more sophisticated. Having a functioning and effective quality assurance process facilitates that products, in this case geodata, meet requirements in relation to cost, quantity, quality and timeliness (ISO, 2012:1,3).
The Swedish government has placed great emphasis on national cooperation between organizations and assigned Lantmäteriet the task of coordinating operations within the geodata sector of Sweden. This responsibility involves establishing and overseeing a national geodata infrastructure in collaboration with local governmental authorities, county boards and other official organizations. The objective of this work is to enhance the access to geodata and interchange of information (Annual report, 2012:28).

In order to establish an effective infrastructure where geodata is made available by different organizations it is necessary to investigate the current geodata quality assurance processes that are present in Swedish organizations. Presently, there exist some procedures for quality check in most organization that handle geodata but in many cases a specific quality assurance process has not been mapped. By analyzing current situations in Swedish organizations, inquiring about their working methods and quality checks it is possible to suggest improvement recommendations and even map a new general quality assurance process for geodata. This could contribute to a more successful cooperation, enhance data quality and help to maintain the established quality level. Consequently, the reliability and validity of available data through the newly established infrastructure would be greatly increased.

1.3 Goal

The goal of this thesis is to have a proper functional quality assurance process so as to eliminate errors and control data processing for keeping the geodata in best quality. Through mapping the quality assurance process, the existing integrated quality assurance process in geodata organizations will be replaced by an enhanced independent one. This will ensure geodata quality in a more effective and efficient manner.

1.4 Scope

The scope is to perform a research on the current geodata quality assurance processes in four different organizations. The research is carried out through analysis and comparison so as to come up with a new improved geodata quality assurance process. It is conducted by interviewing the geodata organizations in order to have a realistic analysis of the procedures of geodata production in terms of; data collection, data analysis, data verification and data entry procedures.

The research is to provide recommendations for enhancing the geodata QA by mapping the current process and the improved one; Is map and Should map. This will contribute to promote the geodata quality.

1.5 Methodology

1.5.1 The aim of the study and for whom it is carried out

The intent of this study is to investigate the quality assurance processes within data management of specific organizations and agencies. This study is being carried out using the knowledge gained and methods learned during the program Project Management and Operational Development at the Royal Institute of Technology (KTH) in collaboration with
Lantmäteriet, a governmental authority responsible to the Ministry of Health and Social Affairs of Sweden. The organization is responsible of supplying society, government, business and individuals, with information about geography and real estate. This is done in cooperation with others in public and private sectors, nationally and internationally. Lantmäteriet recognizes that quality in data has become more and more important, especially since emphasizes has been put on collaboration and interchange of information between authorities and organizations. The available resources for this study are:

- Roland Langhé, supervisor at KTH.
- Jesper Paasch, supervisor at Lantmäteriet in Gävle.
- Other employees involved in the data QA process at the interviewed organizations.
- Process documents and other related files from the interviewed organizations.

1.5.2 Mental models and type of study

Determining the focus of a study is very important. In order to do so successfully, mental models need to be defined. The mental model used in this study is the systems approach model. In order to analyze QA processes one needs to investigate the system as a whole for one factor will most likely affect another. In the systems approach model the reality is mapped objectively and the interdependence between elements investigated.

This thesis is conducted as a qualitative research where it is sought to gain and in-depth understanding of certain phenomena or situations and to discuss peculiarities. Here the emphasis is on investigating the why and how, not just the what, where and when. Furthermore, this research is a case study where existing QA processes in particular organizations are thoroughly analyzed, identified problems discussed and improvements suggested.

Being a case study, the students benefit from descriptive reasoning as well as inductive reasoning. Descriptive reasoning is used to identify pros and cons of current QA processes while inductive reasoning is used to develop a new hypothesis by designing a functioning QA process.

1.5.3 Limitations

During this investigation some limiting factors were encountered. The time constraint for this thesis could be classified as the most limiting factor. The time limit to finish this study was set to 10 weeks. During that time the students were able to conduct satisfactory research of the relevant areas and reach the goals of the thesis within the determined scope. However, given more time the research could be more detailed and cover broader range. Regarding the interviews, the time constraint also limited the number of interviews that could be scheduled and executed. Moreover, it limited the number of data management processes that could be analyzed. Planning the interviews was a time consuming process because of overlapping, vacation periods and other general unavailability of the potential interviewees.

Due to the location of the main organization, Lantmäteriet, which is situated around 180 km north of central Stockholm, the students did not have the opportunity to work at site during
the whole thesis period. If the situation would have been different allowing the students to be in constant contact with all the employees involved in QA at the organization a deeper understanding of the current processes would have been gained. Furthermore, a more comprehensive observation of the working methods and environment could have benefited the study.

1.5.4 Existing knowledge and questions to be answered

Nowadays, extensive amount of knowledge exists about data management and the importance of maintaining quality requirements. Given the large amount of time and effort devoted to data quality throughout the years, one might think that data quality issues have all but been resolved by now. However, a brief look through available literature on the matter reveals that data quality is still an interesting topic and new material and discussions about the subject are being published regularly. Although standardized procedures and general standards for data quality have been developed, every QA process is unique and therefore has to be managed as such.

The two main questions to be answered in this thesis are:

- What is the quality assurance process for geodata at Lantmäteriet, Stockholms Stad, Sjöfartsverket and iMMAP?
- Can the process be improved? If so, how can it be improved?

In order to answer the main questions some sub questions need to be addressed. Those sub questions are for example:

- What is the present situation at above mentioned organizations?
- What disconnections in the current QA process can be identified?
- Can comparison with QA processes in other similar national or international organizations be of aid? If so, what is the possible gain of such comparison?
- What is the root cause of the problem?
- What possible improvement measures can be carried out in order to solve the identified disconnections in the QA processes?

1.5.5 Tools

In discussion of the methodology of this thesis it is important to specify the tools and methods used in the investigation. The main tool employed for data gathering is interviews. The interviews are carried out face to face with predetermined questions, designed to lead the interviewee in the right direction and exploit the time of professionals to the maximum. Using interactive interviews reduces the likelihood of misinterpretation of the questions and ensures the same understanding at the matter at hand. In addition to the interviews, short questionnaires are developed and sent via e-mail to relevant professionals in the field (in most cases previous interviewees) where remaining points and unclear matters from the previous interviews are covered.
Other tools used are literature survey, general observations and brainstorming. The literature survey is a research of the relevant area of the thesis within the specified scope. This area covers general discussion about quality in geographic data, clarifying matters like what is quality and how are organizations ensuring and maintaining it. Analyzing existing knowledge within the area enables students to explore, question and build on that knowledge.

General observations are conducted at site by the students. The interviews are executed at the organizations in question. During the interviews the students are given the possibility to observe to some extent the working habits and sense the atmosphere at the workplace. Brainstorming is a common creativity technique used excessively in this study in order to generate ideas about problem areas. Using this method the students are more likely to be creative in thinking and go out of their comfort zone.
2 Theoretical Reflection

In this chapter the theoretical background for this study will be discussed. The chapter provides definitions of geodata, GIS, quality assurance and quality control. Moreover, it lists potential user groups and application scenarios and discusses uncertainties in geodata and why planning for quality assurance is important. Lastly, principles of data quality, such as vision, policy and strategy, are identified along with a description of systematic quality assurance process management and process mapping.

2.1 Geographic data and GIS

Geodata is basic facts and statistics about topography and environment of a specific area that is in use or will be used for a certain service. Every dataset that has a spatial feature can be defined as geodata; it could be called spatial data, geographic data or GIS data. The prefix geo means that the dataset has a spatial feature or component. The term geodata refers only to data regarding the earth; otherwise when including other planets then the term spatial data should be used (GeoVITe, 2010).

Geographic Information System (GIS) consists of many important components, but geodata is the most important one of the system. Geodata has various usage possibilities, for example it could be used for conducting queries, simulations and spatial analysis. Usually, geographic base data is provided by national authorities or international agencies responsible for surveys and maps and mainly includes topographic information saved in maps or landscape models (GeoVITe, 2010).

Spatial base data and thematic data are subsets of geodata. Aerial and satellite pictures are considered as spatial base data, given that they only present topographic information. On the other hand, thematic data could include a geometry component and is most of the time linked to the spatial base data with using coordinates, admin units, addresses or zip codes as can be seen in Figure 1. The important components of geodata are attribute, time and space, which describe the what, when and where. Figure 2 illustrates geodata important components along with an example for further explanation (GeoVITe, 2010).

Geographical Information System (GIS) is a structured and organized collection of computer hardware, software, geodata and experienced operators. GIS is management, modeling, analysis, simulation and presentation of geodata through a computerized system. In the past, GIS needed very expensive hardware but today’s GIS software functions with standard personal computers. Here it is important to mention that the hardware is the cheapest element of the system, while the software is more expensive. Concluding, the most expensive parts of a GIS are geodata and skilled operators. Figure 3 displays GIS components (GeoVITe, 2010).

The widest use of geodata is for modeling, analysis, simulation and presentation functions. It is assumed that about 80 – 90 percent of digital data has some spatial aspect or can be linked to an existing geodata. The most popular usage groups and scenarios are listed in Table 1 (GeoVITe, 2010).
Figure 1. Typical Thematic Layer Model in GIS (GeoVITe, 2010).

Figure 2. Important Components of Geodata: Space, Time and Topic (GeoVITe, 2010).
Table 1. Popular geodata usage groups and scenarios (GeoVITe, 2010).

<table>
<thead>
<tr>
<th>Usage Groups</th>
<th>Usage Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveying &amp; Photogrammetry</td>
<td>• Newly acquired data verification with existing geodata</td>
</tr>
<tr>
<td></td>
<td>• Merging new data into existing data sources</td>
</tr>
<tr>
<td></td>
<td>• Geodata Visualization</td>
</tr>
<tr>
<td>Cartography</td>
<td>• Producing maps</td>
</tr>
<tr>
<td></td>
<td>• Producing 3D visualizations</td>
</tr>
<tr>
<td></td>
<td>• Analysis of historical maps</td>
</tr>
<tr>
<td>Physical Geography and Geology</td>
<td>• Producing geologic maps</td>
</tr>
<tr>
<td></td>
<td>• Analysis of the potential of natural hazards</td>
</tr>
<tr>
<td></td>
<td>• Terrain surfaces modeling, geologic structures and climatic modeling</td>
</tr>
<tr>
<td></td>
<td>• Terrain surfaces 3D visualization</td>
</tr>
<tr>
<td></td>
<td>• Analysis, visualizing and modeling of hydrologic systems</td>
</tr>
<tr>
<td>Human Geography</td>
<td>• Producing thematic maps</td>
</tr>
<tr>
<td></td>
<td>• Analysis and modeling the socio-economic phenomena</td>
</tr>
<tr>
<td></td>
<td>• Geostatistics</td>
</tr>
<tr>
<td>Leisure Activities</td>
<td>• Routing and navigation services</td>
</tr>
<tr>
<td></td>
<td>• Leisure time activities, planning and documenting</td>
</tr>
<tr>
<td>Telecommunication, Supply &amp; Disposal Industry</td>
<td>• Network systems planning and maintenance</td>
</tr>
<tr>
<td>Medicine</td>
<td>• Analysis of spatial distribution and spreading of diseases</td>
</tr>
<tr>
<td>Botany and Zoology</td>
<td>• Plants and animals studies</td>
</tr>
<tr>
<td>Marketing and Financial Services</td>
<td>• Geo-Marketing</td>
</tr>
<tr>
<td></td>
<td>• Optimizing potential store locations</td>
</tr>
<tr>
<td></td>
<td>• Real estate business</td>
</tr>
<tr>
<td>Logistics</td>
<td>• Fleet management</td>
</tr>
<tr>
<td></td>
<td>• Route optimization</td>
</tr>
<tr>
<td></td>
<td>• Vehicle navigation systems</td>
</tr>
<tr>
<td>Urban &amp; Regional Planning</td>
<td>• Analysis of socio-economic phenomena and patterns</td>
</tr>
<tr>
<td></td>
<td>• Modeling impacts of political decisions with a spatial aspect</td>
</tr>
<tr>
<td></td>
<td>• Visualizing planned changes in landscape and cityscape</td>
</tr>
</tbody>
</table>

Figure 3. Components of GIS (GeoVITe, 2010).
For more comprehensive discussion about geodata and GIS, benefits, purpose, formats and etc., readers are referred to Appendix A.

2.2 Quality Control and Quality Assurance

Before venturing into quality management in geodata it is important to define few concepts first such as quality control and assurance. Quality control (QC) and quality assurance (QA) consist of activities to ensure the quality of a particular result, often referred to as a service or a product but in our case geodata. These topics, QC and QA, are usually used interchangeably but their focus is completely different. Quality control is a reactive process that focuses on identifying defects and errors while quality assurance is a proactive approach which purpose is prevention (NBDPN, 2004:7/6-7/7).

In QC observation methods such as checking, investigating and discovering are used to find inaccuracies, defects and areas where requirements for quality are not fulfilled in the final outcome. Results from the QC are used to contain, evaluate, adjust and resolve the defects found in order to enhance the accuracy of the geodata. In QA systematic activities are planned and implemented in the quality system processes to hinder inaccurate and imperfect development of data. Therefore, it can be stated that QC often results in QA, where QC identifies defects and QA remodels the process to eliminate the defects and prevent them from recurring. Furthermore, it can be claimed that one cannot exist without the other if data quality objectives are to be met (NBDPN, 2004:7/6-7/7 and Chapman, 2005:5).

QA is the main pillar in any successful data gathering. If quality is lacking in the data it has lost its credibility and can therefore be deemed expendable. Planning for QA in geodata is very important. Most people recognize the feeling of a very busy Sunday where you wind up spinning your wheels, accomplishing nothing instead of achieving all your tasks. Without planning the work you are endlessly retracing your steps, performing the same work multiple times and often in illogical order, having difficulty remembering what to do next and making decisions in a hurry (EPA, 2003:2-3).

2.3 Planning for QA in Geodata

If a busy Sunday can easily go astray, an unplanned quality management system will most definitely fail. A plan certainly defines what is expected of the QA and success is almost unattainable if that definition has not been clearly formed and agreed upon. QA planning in geodata may save a great deal of time and effort seeing that concerning parties are better suited to identify potential problems that may affect data quality, budget or schedule. Generally, with no planning projects tend to cost more and take longer time. Concerning geodata, if QA planning is rushed or avoided altogether there is a high probability that the data needs to be corrected or redone, that is gathered, analyzed, processed and represented again. Investing time and money in proper QA planning will pay off in the long run since inadequate planning can lead to poor decisions resulting in budget and schedule overrun that no one could have anticipated (EPA, 2003:2-3).

In continuation of this discussion, a plan for QA and QC should be developed parallel with establishing criteria for data quality acceptance. Furthermore, the testing categories and
requirements should be defined (Intetics, n.d.). The purpose of a QA plan is extensive. A QA plan documents the results of a systematic planning process and provides a complete, clear and concise schema for geodata operations. Furthermore, it provides the quality objectives, the data collection and processing methods, the assessment procedures to confirm if the quality of the output data is sufficient, publishing methods and any limitations on the use of the output data. The extent of detail for QA plans is dependent on the type of data to be acquired and processed, questions to be answered and decisions to be made (EPA, 2003:7 and EPA, 2002:4). Usually, a QA plan includes four main parts (Intetics, n.d.):

- **Organization part**: The scope of the QA process is outlined along with the QA project background to ensure the correct understanding of the process. Here, the QA team is structured and described in terms of roles and responsibilities.
- **QA test design and process**: The entire testing process is outlined with defining a list of required tests. Also, the acceptance criteria are defined during this phase.
- **Quality assessment**: Every detail of each test is provided by specifying types and methods of quality checks. Additionally, tools for automated testing are determined.
- **Quality records**: A detailed quality report is produced on the results of each test conducted and analyzed.

In regards of quality checks, they generally consist of the following steps (Intetics, n.d.):

- Initial checks; batch checks, using scripts or automatization tools
- Topology checks
- Visual review
- Quality control report

### 2.4 Data Quality and Uncertainties in Geographic Data

Defining, measuring and estimating quality allows us to compare data and information and conclude which data is the best and the worst. Concerning geodata, quality is usually thought of as the degree to which particular data is fit for certain application, or “fitness for purpose”. Satisfactory quality depends on the application in question but the guiding principle is estimating how much uncertainty exists in the data and deciding how much uncertainty is acceptable. In all geodata some degree of uncertainty is always present. Perfect data is impossible to create and even if we possessed such data it would be too large, detailed and expensive for pragmatic usage (DiBiase et al.). Other essential factors about geodata quality can be seen in Table 3 in Appendix A.

When assuring quality in geodata it is important to identify, quantify, track, reduce, report and represent uncertainties. Since our nature and landscape is continuously changing it is very difficult to obtain accurate geodata. Furthermore, budgets, measuring equipment constraints and human capabilities add to the uncertainty and the geodata produced is merely approximation of the reality. The difference, or uncertainty, between the data and the reality propagates, and is often magnified, through data processing such as data collection, analyses and representation (UCGIS, 2007).
Geodata consists of typology (the type of geographic feature), location and spatial dependence (the closeness to other geographical features) which all involve uncertainty. In addition, uncertainty lies in every phase of geodata life cycle; data collecting, analyzing, processing, evaluating, representing and final results. In order to improve quality in the data and minimize factors that invite inaccuracies following tasks could be carried out (UCGIS, 2007):

- Examine how uncertainty arises and propagates through geodata life cycle.
- Design methods for minimizing, quantifying and representing uncertainty and how its propagation can be predicted.
- Identify and develop new activities for uncertainty management in geodata.
- Implement strategies and policies about activities in every phase of the data life cycle such as gathering, quantifying, evaluating and documenting.
- Enhance uncertainty documentation throughout the geographic life cycle.
- Improve communication about uncertainty discoveries.

People often act under the assumption that geodata is free of errors because it is produced by a computer. High confidence in software computation is the result of continuous development in the technical industry. Although models today are accurate and produce precise outcomes these results are dependent on the accuracy of the input data. A single input data can affect multiple outcomes and whether it is acknowledged or not there is always some extent of uncertainty involved in every data gathered. Data analysis operations alter this uncertainty even further, usually in the form of amplification. It cannot be objected that computers are good at computations and rarely make mistakes. However, our computer calculated results will only be as accurate as the input data and where there is human involvement errors are inevitable (Keukelaar et al., 2000).

2.5 Principles of Data Quality

Having efficient QA in our data management processes enables us to treat data as a long-term asset, generating future growing value for the organization. In QA geodata quality is improved by focusing on prevention and correction. Greater emphasis is placed on prevention since it is both more time consuming and costly to identify and correct errors. However, data correction is a very important factor in data quality improvement since it is almost impossible to prevent all errors to emerge. In QA data quality principles need to be carried out in every part of the data management process; collection, entry, storage, analysis, publication and usage (Chapman, 2005:8). Establishing a geodata quality policy and strategy forms the base of these principles. Chapman (2005) formulated this nicely in his report for the Global Biodiversity Information Facility:

   Begin by setting a data vision, developing a data policy and implementing a data strategy – not by carrying out unplanned, uncoordinated and non-systematic “data cleaning” activities (Chapman, 2005:8).

In order to have quality in geodata the organization needs a vision supporting its quality goals. A data quality vision unifies organizational actions towards data quality, describes key values
for long term data and informational needs and acknowledges data and information as the most fundamental part of the organization (Chapman, 2005:8).

A data quality policy must accompany the vision in order for successful implementation and maintenance of QA in geodata. A data quality policy defines the organization’s data management processes, clarifies goals with regarding data quality improvements, enhances communication with and between data providers and users and reinforces the credibility of the organization and the quality of its data (Chapman, 2005:8-9).

Considering that large organizations like Lantmäteriet possess and control extensive amount of geodata there is a necessity for creating data quality strategy for collecting and analyzing data. Developing a data quality strategy one needs to bear in mind several factors (Dravis, 2004):

- **Context**: What type of data is being improved and its purpose of use.
- **Storage**: Where and how data is being stored.
- **Data flow**: How data is entered and its flow through different parts of the organization.
- **Work flow**: How data management processes interact with work activities.
- **Conservation**: The people responsible for managing the data.
- **Monitoring**: How data is continuously checked and updated.

Goals are what drive strategies. The above mentioned factors are centered on the goals of the data quality management processes which can be short, intermediate or long term. Good data management principles such as reducing duplication, motivating sharing of tools and information and using standards as much as possible should be included in the quality strategy (Dravis, 2004 and Chapman, 2005:9). Standards ensure a common understanding of what quality is and how it can be expressed and measured. Numerous standards exist related to quality and quality management and even more related to geographic information. The International Organization for Standardization (ISO) develops large amounts of international standards that range from food safety to computers, and agriculture to healthcare. In Table 2 a few ISO standards are enumerated and described briefly. The listed standards mostly discuss general quality management and management related to quality in geographic information which is the focus in this study. Indeed there exist other organizations that develop standards and frameworks for quality and geodata but discussing those in detail is beyond the scope of this thesis.
Table 2. Quality and GEO information related ISO standards. Information retrieved from ISO’s (International Organization for Standardization) official website www.iso.org.

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Short description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISO 9000</td>
<td>Quality management systems: Fundamentals and vocabulary</td>
<td>Specifies essentials of quality management systems and lays the basis for the ISO 9000 family.</td>
</tr>
<tr>
<td>ISO 9001</td>
<td>Quality management systems: Requirements</td>
<td>Describes requirements for quality management systems, focusing on meeting customer needs and increasing customer satisfaction.</td>
</tr>
<tr>
<td>ISO 9004</td>
<td>Managing for the sustained success of an organization: A quality management approach</td>
<td>Guidelines for performance improvements and sustaining success by a quality management approach</td>
</tr>
<tr>
<td>ISO 19113</td>
<td>Geographic information: Quality principles</td>
<td>Lays foundation for describing and communicating quality in geographic data.</td>
</tr>
<tr>
<td>ISO 19114</td>
<td>Geographic information: Quality evaluation procedures</td>
<td>Describes procedures for determining and evaluating data quality, applicable to what is defined in ISO 19113</td>
</tr>
<tr>
<td>ISO 19115</td>
<td>Geographic information: Metadata</td>
<td>Delineates a schema needed to describe geographic data (identification, magnitude, quality, spatial reference, distribution).</td>
</tr>
<tr>
<td>ISO 19131</td>
<td>Geographic information: Data product specifications</td>
<td>Provisions for specification of geographic data products, formed on principles from other ISO 19100 standards.</td>
</tr>
<tr>
<td>ISO 19138</td>
<td>Geographic information: Data quality measures</td>
<td>Describes data quality measures. The selection of a measure is dependent on the type of data and purpose of use.</td>
</tr>
<tr>
<td>ISO 19158</td>
<td>Quality assurance of data supply</td>
<td>Presents foundation for quality assurance specifically for geographical information. It is based upon quality principles defined in ISO 19157 (under development) and ISO 9000.</td>
</tr>
</tbody>
</table>

2.6 Systematic QA Process Management and Process Mapping

The aim is to have a clear and direct control over the QA process, so as to enable a smooth improvement and tracking paths. Without a process map, the management of the process will be difficult especially in identifying the defects and disconnections that will be involved in the improvement of the process (CDF, n.d.).

For a better control and management in a QA process, like in any other process, there should be a process owner and a policy. The head of the organization is responsible for the overall QA, through the middle managers and all involved staff in the process, and is committed to support the QA management team at all times. It is the responsibility of the QA section in the organization to hold meetings regularly to assess and evaluate the QA process and working methods in order to review and determine the essential areas for improvement. Also, the QA section is responsible to the head of the organization for reporting the quality status and improvement recommendations. The entire organization management team is responsible for delivery of statements confirming the quality assurance and improvement of the work practices and activities (CDF, n.d.).
The QA policy is to reflect an effective and efficient functionality of the process and the responsibility for implementing it. It requires commitment and ownership from all staff across the organization, but will be under control of an assigned section or unit for easier defect tracking and finding areas for improvement. Taking the quality assurance into high consideration in geodata organizations, when undertaking their work, is vital. To that effect and according to the QA policy, the organizations shall (CDF, n.d.):

- Sustain the work consistency methods in accordance to established policies, procedures and regulations, so as to avoid deviations from the scope and strategies.
- Make sure that policies, procedures and regulations are executed and reviewed systematically.
- Monitor and measure regularly the quality of geodata processing methods in order to achieve high quality of geodata outputs (for publishing) and maintain the best value with continuous enhancement.

This will be achieved when a clear structured QA process map is available based on a well established policy (CDF, n.d.).

In order to be able to manage the QA process successfully, systemizing the management of the process is one of the best solutions (Rummler and Brache, 1995:164-168). For an efficient customer orientated management system, the top management should establish the organization accordingly. As mentioned above, QA process mapping will ensure that. Without a process map, it is hard to measure the time that the sub-processes and the total process would take, and also it is difficult to find the disconnections in the process. The QA process map could be a great help in using the Lean Production philosophy in the organization for utmost performance. So, the QA process map is the most important support process that sustains the entire geodata management system and avoids deviations from the scope (Langhé, 2013).

‘Is’ and ‘Should’ mapping is the main step in process improvement projects. An ‘Is’ process map represents the current situation in an organization. It shows the input-output relationships among departments as well as illustrates the steps that departments carry out in order to convert inputs to outputs for a particular process. Mapping the ‘Is’ process allows for critical interfaces, overlays and disconnections to be identified. Once disconnections have been identified, a ‘Should’ map can be created, where improvement measures have been implemented, reflecting the desired disconnect-free process (Rummler and Brache, 1995:49).
3 Fact Gathering

The fact gathering for this study was an ongoing process throughout the entire thesis work. The main data collection was carried out through on-site interactive interviews. Three Swedish organizations, Lantmäteriet, Stockholms Stad and Sjöfartsverket, were interviewed for the thesis. These organizations play a big role in the Swedish geodata sector and could be classified as frontrunners in terms of collecting, analyzing and publishing geographical information. In addition to the Swedish organizations, inquiries were made to an international nongovernmental organization, iMMAP, about their ways of conducting quality assurance in data management.

In each interview several key persons were present. These employees were the ones most involved in the quality assurance process at their organization and accustomed in managing geodata quality issues and maintaining required level of quality in their data management system.

The interviews were structured in the following manner:

- A short introduction about the organization in question and its operations, given by an employee at the organization.
- An allocated time slot where students were given the opportunity to ask predetermined questions.
- A final discussion between the students and interviewees where additional topics and questions that emerge during the interview were addressed.

During the interviews the students got an overview of the current quality assurance processes and procedures at the three organizations listed above. Furthermore, the organizations’ specialty, goals, future work, policy, mission and strategy were presented to the students. The design of the interviews allowed the interviewees to express their views on the current situation in their company, identify potential improvement areas and suggest improvement measures they felt to be suitable for their organization.

In addition to the interviews a significant part of the data was collected through personal communication, mostly online through e-mails or face to face short informal meetings. Through these communication channels hard copy documents, especially from Lantmäteriet, were gathered discussing areas of QA and illustrating current QA processes at the organization. For some documents matters of confidentiality needed to be considered. In those cases the students strove to comply with the wishes of the concerned organization.

Due to the nature of the organizations most documents that the students received were in Swedish. Moreover, some parts of the interviews were carried out in Swedish but English used to explain complex and confusing matters in more detail. Although both students have a decent background in the local language their nationality is not Swedish. Because of that the students had to overcome some difficulties of language barriers. However, due to the diversity of the partnership one student’s weakness could be compensated with the strength of the other and vice versa.
In terms of the literature survey, theoretical data was obtained from academic books, published articles and journals, the internet and course material from the program Project Management and Operational Development at KTH. Through surveying available material, existing knowledge on the particular subject was explored in order to prepare the students for the work ahead.

Regarding the reliability of the data, this study is based on research technique and qualitative analysis. Data collection is mostly through interviews where answers can be subjective and individual personal perspectives potentially influence the outcome of the study. However, in order to keep the data reliability high it is sought to rely on physical documents where possible and refrained from being affected by personal views. For the theoretical data, the reliability can be regarded as much higher for in that case ideas are obtained from previously published and authenticated data.
4 Description of Fieldwork

Field work activities were carried out at Lantmäteriet office in Gävle, Stockholms Stad office in Stockholm and Sjöfartsverket office in Norrköping. The field work was divided into three main phases; understanding the current quality assurance process, scrutinizing the current process to find disconnections and design solutions to create suggestions for enhancement of the current process. These phases were completed through interviews, meetings and email correspondence where different information in document, open discussion and presentation forms were provided by the mentioned organizations and then studied for the thesis report.

As the original party for this study is Lantmäteriet, the field work began there by a preliminary meeting with the supervisor. During that meeting a background of the organization was presented with a brief discussion on the thesis subject; Quality Assurance in Geodata. The meeting continued with an open style presentation by quality management responsible and discussions through a list of questions prepared previously for this purpose. The overall meeting provided thorough information on the entire organization management system and different roles in the data management process which enabled the students to understand the current QA process in the organization. In continuation of the meeting, through direct and continuous communication via email correspondence, a thorough study was conducted on the current QA process map.

For field activities at Stockholms Stad, the information gathering on their QA procedures and geodata management was through a meeting held at their main office in Stockholm. In this meeting, where related personnel were assigned to attend and contribute to the presentations and discussions, detailed information was received based on the list of questions sent earlier by the students. The presentation was very thorough and the QA topic was covered very well. In order to practice a better understanding of Stockholms Stad’s QA process a direct study was conducted on the findings where the QA process was described for further comparison with Lantmäteriet’s QA process.

As for Lantmäteriet and Stockholms Stad, the field work with Sjöfartsverket was started by direct communication with their appointed staff for the thesis interviews and sending the list of questions in order to have a clear view of what would be needed during the interview. Then through a very responsive coordinated interview at their office the requested list of questions from the students was answered by the assigned staff in an open discussion session. A comprehensive presentation was conducted during the meeting, and involving detailed information about their entire geodata processing and QA.

Because of the work experience and having communication network internationally, the students contacted a humanitarian Non-Governmental Organization (NGO) involved in information management of mine action, called Vietnam Veteran of America Foundation’s Information Management and Mine Action Programs (VVAF iMMAP). This was done in order to study their data quality management and QA process. The process map analyzed was their Information Management System for Mine Action (IMSMA), which involves GIS for producing maps.
After performing the above mentioned part of the field work, that is scrutinizing the main data management system of the organizations, an in-depth description of the current QA processes was made. That was done in order to identify defects and areas of enhancement, find solutions and suggest a more proper way of managing the quality assurance process in form of a ‘Should Process Map’.

The entire field work was done through direct advices from the KTH-examiner and Lantmäteriet supervisor. Most of the research and studying activities of the current processes were done at KTH facilities in a continuous commitment with daily activity plans.

Firstly, a tentative time plan was developed, and accepted by KTH-examiner, in order to have a clear schedule for progress. Further to this, a preparatory meeting was held with Lantmäteriet supervisor so as to plan for the field activities, trips, meetings and interviews. This involved finding other interesting geodata organizations, for studying and comparison purposes. Direct communication was practiced by Lantmäteriet supervisor, which facilitated the other parts of the field work that involved both of Stockholms Stad and Sjöfartsverket, and encouraged the students to have a greater motive to conduct the field activities and study sessions.
5 Analysis of Data and Information

In this chapter a background of all involved geodata organizations, including their main roles and responsibilities within the data management, is introduced. Furthermore a detailed description of their general management system and an analysis of their respective current QA process is included, followed by QA process maps if available. The information in this chapter is based on the interviews performed at each of the organizations (excluding iMMAP, which was done through personal communication). The interview questions can be seen Appendix B. The same list of questions was used for Lantmäteriet, Stockholms Stad and Sjöfartsverket.

5.1 Lantmäteriet

5.1.1 Introduction

Lantmäteriet carries out their activities on behalf of the Swedish government and is responsible to the Ministry of Health and Social Affairs. The organization employs over 2,000 people situated at 70 office locations all over Sweden. It is organized in four divisions and a number of general support functions such as IT solutions and maintenance. Lantmäteriet divisions are (Lantmäteriet Annual Report, 2012:31):

- Cadastral services: Responsible for property division, i.e. deciding on new properties or modification of existing boundaries.
- Property registration: Decides on and registers ownerships, leaseholds, mortgages and other related issues.
- Information development: Develops and provides geographic and real property information.
- Authority missions: Works with military applications.

Lantmäteriet’s mission is to contribute to social and economical development by creating the conditions for forming and developing real property and national infrastructure, buying, owning and selling real estates and seeking, finding and using geographic and real property information (Lantmäteriet Annual Report, 2012).

Lantmäteriet has been mapping Sweden since 1628. Although it is one of the oldest organization in Sweden Lantmäteriet considers itself as a modern organization that evolves over time in order to meet future requirements. Lantmäteriet does not only develop maps and demarcate land and property boundaries (Lantmäteriet Annual Report, 2012:31). They for example receive large amount of data from municipalities and local governmental authorities about addresses, selling and buying records, ownership, building types and future plans, that they need to manage, adjust, organize and publish. For more elaboration on Lantmäteriet operations see Appendix C.

5.1.2 Analysis of the current QA process in Lantmäteriet

Lantmäteriet considers quality assurance as a crucial aspect of the data management system, in order to achieve the required data specification, level of accuracy and reliability. As it is the
responsible authority of surveying, managing geodata and mapping for the whole country, the quality of the geodata is taken as an important process in the GIS. It was seen necessary to have data quality assurance process, specifically, in two main stages; before data storing and before data publishing. This is believed to promote data reliability and maintain the data quality; in turn meeting data users’ needs and expectations.

Lantmäteriet is an organization with numerous employees scattered all over Sweden. It has developed through its existence and will continue to do so in the near future with new activities and responsibilities merging into the organization and becoming a part of its operations. Due to its size, number of employees and continual growth and development it has proven difficult to develop an integrated organizational culture. Although Lantmäteriet suffers from subcultures within the organization the general opinion is that data quality, its importance and usefulness, is incorporated into the organizational culture or cultures. Seen as a part of the culture the importance of quality is communicated through the organization with weekly meetings for the information management group where imminent quality problems are discussed, solutions suggested and decisions made.

When inquiring about the organization’s policies and strategies it became clear that Lantmäteriet has no specific policy regarding the quality of its operations and production. Further discussion about the topic revealed that our interviewees believe that the importance of quality is integrated into the overall policy and strategic objectives of the organization. However, after some investigation it did not become clear that data quality was given significant weight in the organization’s mission or strategic plan although it is mentioned in couple of places.

Despite the fact that a specific quality vision, policy and strategy is mostly absent from their current data management Lantmäteriet takes deep interest in the quality of the data they produce. That can be seen through their actions towards implementing international standards for maintaining the quality of geographical information. Currently, Lantmäteriet is not actively using the ISO 9000 series, although some parts are ISO 9001 certified, but focuses instead on standards that are directed specifically at geographical information. They are using ISO 19113, 19114, 19131 and 19138 to evaluate a limited number of quality themes e.g. completeness, and ISO 19115 for metadata at the National Geodata Portal (see description in Appendix C). In the near future, ISO 19157 will replace ISO 19113, 19114 and 19138 but it is still under development. Other standards that Lantmäteriet have been considering are for example the newly released ISO 19158 (a more detailed description of the ISO standards can be seen in Table 2 in the chapter 2.5).

In Lantmäteriet, there is a functional QA process that is integrated into the entire geodata management system and seen as an important aspect in the organization culture. Although the QA process for the entire organization is not designed as a separate process and to be implemented by a specifically skilled staff of a department in the organization, it is functioning well and figured out as a simple and inexpensive process to implement. Still, there is emphasis for improving the current QA process, as well as setting specific quality goals.
The QA procedures truly involve every employee in one way or another for ensuring the required level of quality acceptance. Take for example employees in data management. Lantmäteriet receives extensive amount of data from numerous municipalities in Sweden. Once a delivery from a municipality has been acquired the same employee is responsible for following the data delivery through the entire data management process. This data management process consists of receiving a data delivery or requesting for one, inspecting and analyzing the data to ensure it fulfills quality requirements, updating and publishing the data and contacting municipalities if defects are found that need adjusting. It can be said in this context that everybody is responsible for “their” data or rather “their” municipality.

The close collaboration between the municipalities and the data collection section in Lantmäteriet motivates the data source to be familiar with the data specification requirements and send in required data quality. According to the agreement between the municipalities and Lantmäteriet, data provision and reception is taking place. In order to make sure that the agreement between the two parties is fulfilled relatively, there are regular annual meetings between them. After validating the data by the data collection section, if any errors are found, it is the responsibility of the data provider to rectify those errors.

Lantmäteriet receives data from municipalities and collects data such as through aerial photography and laser scanning. There is no difference between the data quality check in both self data collection and outsourced data collection processes. For example, when collecting address data, the municipalities are solely providing those data according to the mentioned agreement. On the other hand, regarding building data collection, Lantmäteriet and municipalities are collaborating based on the same guidelines; Lantmäteriet controls via aerial photography, whereas the municipalities control the actual building. This can lead to different interpretations but more often than not the municipalities can correct Lantmäteriet’s interpretation.

A manual method is used in the data quality QA in the form of an error documenting system designed for managing the quality assurance process. An outsider would need a thorough look to understand the system, which could take longer time than a person familiar with the process would need. For indicating errors, Lantmäteriet uses flags like “the building is not located on an address”. When errors are detected through the filtering procedures, reports are created describing the errors and the correction actions needed. These error reports, including recommendations for solving errors, will be followed up by giving feedback to the municipalities, which is very important to avoid further issues. Furthermore, there is a possibility to make changes to operation systems if that would mean preventing repeating errors.

Based on the origin, size and type of the error or issue (if the error is repeatable or not), it will be reported to relative sections, such as the Base Data Group (Grunddatagrupp), for corrections and solving. Small errors are usually dealt with and solved right away but larger ones are be reported by the Information Developer for further solution. Solving those errors is conducted by the Information Management or System Managers. Some errors might only need better or more comprehensive handbook and clearer guidelines to be solved, but others
need further investigation. Lantmäteriet provides the municipalities with clear specifications of the data they request or receive according to the agreement. These specifications serve the purpose of solving potential quality issues by preventing errors before the municipalities send in their data.

In some cases, data is put through Lantmäteriet’s filtering system even though it is lacking required quality. These cases typically involve data that fulfill the basic specification requirements but some details are missing or something needs updating. It is believed that it is better to publish the defective data if it serves its purpose of presenting valuable information rather than chase after insignificant errors which can often be expensive and time consuming. Another reason for publishing defective data is that it is often known that the data will be updated, anyhow, no later than the next year.

Lantmäteriet has not been actively comparing their data quality assurance process to any other related agencies or organizations in order to benchmark the quality assurance. The organization’s operations are quite unique and benchmarking has not really been discussed as an improvement measure for the organization. However, our interviewees believe that such a comparison could benefit and improve the current QA process in such a way that the process could be described more uniformly and more efficient methods for QA developed. Additionally, this kind of comparison can be particularly justified since the demand for collaborating, information sharing, enhancing access to geographical data and making it available is continually increasing.

5.1.3 Process maps

Monitoring failures during geodata processing is the responsibility of all involved staff; no specific personnel are assigned for conducting this monitoring process. In case of finding a data quality failure, it is reported to related section for treatment and correction. After the failure is registered, an analysis is conducted so as to identify the type of the failure and its level of effect on the final data product. Then a decision is made on the action for correcting the failure and controlling it in order to finish the data processing and have a ready product for delivery. The following process map, in Figure 4, is a simple model of monitoring failure cases.

![Figure 4. A simple process map for monitoring failures at Lantmäteriet (Paasch, J., personal communication, 10th April 2013).](image_url)

In Lantmäteriet there exist many process maps for different activities within the organization. The main process map for the information management can be seen in Figure 12 in Appendix D. The information management system consists of three processes, two main ones and a support process. The two main processes are control and running processes, respectively. In
running process the main data management activities take place such as data collection, storing and delivering.

As described before the data quality assurance process is integrated into the entire data management process. This is clearly illustrated in Figure 13 in Appendix D which shows the process for receiving, verifying and updating data about addresses, buildings and topography (ABT). This process is very detailed consisting of many steps. Some are easily understood and related to but others are more specific to Lantmäteriet’s operations, causing external observers to apply some additional speculations.

Another process map for data management available is a support process for buildings, addresses and apartments (BAL). The process map can be seen in Figure 14 in Appendix D. It illustrates a case management process for the BAL section of Lantmäteriet. The whole process goes through quality check, analysis, investigation and feedback phases. The process ends with correction and completion of each case by reconnecting to the municipalities through regular communication.

The above mentioned process maps are only fraction of what Lantmäteriet are using in data management. The process maps in this thesis serve the purpose of showing how comprehensive the organization’s operations really are. They also show that the entire data management system is in the form of integrated approach in terms of quality management.

As for the IS map, Figure 5 illustrates what Lantmäteriet believes is the current process map for data QA at the organization. Although the main quality QA for the entire organization functions is not mapped separately, but integrated into the entire data management system, it can be expressed for a specific data group or a section management. For example, the following process map in Figure 5 is a sample of managing quality failure, which is designed for quality assurance of data related to the store in the buildings, addresses and apartments section (BAL).

The whole process consists of analysis and management phases. Firstly the quality failure is identified, then analysis and categorization of the quality failures is conducted, and finally the quality failure is managed before data storing for publishing.

The process begins with the sub-process of analyzing and managing quality failure, so as to determine the way of managing the quality failures for supporting customers and the in basic data product. Here, according to the strategy and development plan, quality defects are identified in the data production process in order to manage them through analysis, categorization and definition of the quality failures. In some cases the quality failures are also identified from earlier projects.

Another sub-process, in the quality failure management process, is data error management that results in supplying and delivering the products. This starts by classifying different data source inputs, when collecting data, for insertion. The income data is controlled according to already set up operations rules, formats and regulations in order to be ready for storing in
Develop information and geodata product

Identify quality defects to manage

Need to identify

Analyze and manage quality failure

Identify QF "identify error or failure"
- Logical consistency from specifications
- Matching layer

Analyze QF
- Categorize
  - No action
  - Investigating cases
  - To collection
  - To develop

Manage QF
- Management depends on category / scope
  - Municipality, IN
  - Minor correction, ILL
  - BAL system management
  - Develop
  - No action
  - Information Development

Analyze and manage quality failure

Managed

Customer support (Manage error in basic data product)

An example can be needed

GDS BAL (Basic data store)
LevFR (Delivery data store)
Sigma fastighet (Delivery data store)
Future integration store

Figure 5. The IS map for analyzing and managing quality failures in BAL at Lantmäteriet (Paasch, J., personal communication, 25th April 2013).
basic data store before publishing. In case of detecting errors in the income data after data insertion, before storing in basic data store, the errors will be acknowledged and referred back to the insertion and collecting stage for rectifying the data to clear them for proceeding to store and publish. After publishing the data, there are other activities such as data quality check and flagging what should be stored in the delivery data store. So, from the discussion above it should be clear that there is a continuous data quality check and control activities throughout the entire management of quality failures in the data storing process at Lantmäteriet.

5.2 Stockholms Stad

5.2.1 Introduction

Stockholms Stad or the City of Stockholm has a wide range of activities and responsibilities. The organization is composed of many administrations which are responsible for activities regarding the whole city, such as schools, sports, libraries, road maintenance, real property management, urban planning and the environment. Each administration or department has a board of politicians which have the ultimate authority regarding their activities (Stockholms Stad, 2013).

The topic of this study relates to the activities of the department of City Surveying (Stadsbyggnadskontoret). The City of Stockholm is continuously growing and therefore the need of a thorough plan for the city is increasing as well. The future image of the city involves connecting different parts of the city and blurring boundaries between the inner city and suburban by ensuring high quality architecture, green parks and open spaces. The department of City Surveying is responsible for seeing the future image become a reality. Their roles are vast, including (Stockholms Stad, 2013):

- General and detailed city planning.
- Designing residential and construction projects.
- Strategic planning for water, noise and climate concerns.
- Awarding building permits.
- Producing maps and managing property issues.

Among all the maps they produce of Stockholm, the Base Map is the most detailed one covering the whole municipality. The Base Map is often used for early stage planning and as a basis for other types of maps, such as the City Map and the County Map.

5.2.2 Analysis of the current QA process in Stockholms Stad

In Stockholms Stad the quality assurance of the data is the responsibility of everyone. The ones carrying out the measurements and other data collection bear great responsibility in making sure the parameters are measured with care, the right measurement methods are chosen and the uncertainty is minimized. Furthermore, they are accountable for correcting, updating and inserting the data into the database.
The importance of data quality is being communicated in several ways in the organization. The production groups work according to the lean production practices, where the focus is on cutting down waste while ensuring quality. Furthermore, weekly meetings are held where the production groups discuss quality errors, areas that need improvements and improvement suggestions. Often at the end of the meetings someone is assigned the task of looking into the suggestions and checking out their feasibility. In continuation of this, Stockholms Stad establishes quality projects as required for example when new instruments are being implemented in order enhance the security feeling of employees and guide them in the choice of the right measurement method.

Stockholms Stad believes their current quality assurance procedures are effective but also share the view that it can be improved. As an example, a more comprehensive automatic quality check or assurance is essential in order to prevent unnecessary errors when inserting data into the database. Today, there exists a flagging system for some parameters but others have no restrictions like height values of buildings. In that case it is possible to insert values that exceed the acceptable range without any notification or hindrance. Stockholms Stad wishes to prevail the automatic QA tool across the entire geodata management system, as this will reduce field quality check time and efforts.

Most of the quality check is done at site when measuring and once data has been inserted into the database not many other quality assurance procedures are carried out. This could also be an area for potential improvements where quality check could be established for each measured point in the database.

The flagging system is also used for indicating buildings with permits and locations without permits. This is a good tool for planning purposes and providing strategic data to customers and investment firms, as the flagged addresses without permits are possible locations for new projects and buildings.

For labeling the data there is a software tool in the database that displays the level of accuracy of the data. The boundaries of the addressed locations are colored coded according to the level of data accuracy based on the method used for the data collection. Different methods have different level of accuracies, such as; digital photogrammetry, laser scanning and terrestrial measuring methods. Along with the accuracy it is possible to see the employee responsible for the measurements, the date of the latest update and the measuring method used.

Stockholms Stad is using HMK (Handbok i mät och kartfrågor) for the quality assurance process, which is the Swedish guide book for survey and map issues, developed by Lantmäteriet and based on international geographic data quality assurance standard (Lantmäteriet, 2013b).

Comparisons to other organizations directed at data quality assurance processes have not been actively carried out at Stockholms Stad. However, the organization has been performing benchmarking with other cities around Stockholm, such as Malmö and Gothenburg, in relation with customer satisfaction, working methods and other internal operations. These
cities are organized in different ways but they all have similar issues that can easily be solved by cooperation and interchange of information.

5.2.3 Process maps

The data flow in Stockholms Stad goes through different stages till the data is produced. The production environment data flow is a part of the main data flow. The main clients of this production system are; Web Digital Photo (DP)/Map, Microstation V8 and the Local Road Network. Microstation exchanges data with both Web-server (SBK) and Oracle DP/Spatial database, which share the same data but the data is used for different purposes. Web Digital Photo (DP)/Map only exchanges data with the Web-server (SBK). Cadastral mapping is done in web application, where the system delivers property boundaries via national land survey. Stockholms Stad exports data, from this database, to different databases.

Another database is managed through Traffic Administration (Vägverket), which is Oracle Local Road Network. The Traffic Admin is responsible for the road network in Stockholm and they have their own database called Local Road Network Database (LV). Here, Stockholms Stad is responsible for the geometry in the database and ensuring that quality data is updated and delivered to the national road database.

The databases consist mostly of objects. An object represents different things, for example a building, and can have different geometries for different purposes. These purposes can be geometry for the Base Map, 3D applications or the generalized overview map but all these geometries belong to the same object. Thus, when a building is taken away from the real world, it is deleted in the Base Map, City Map and every other product involving the object.

When an object is changed in a map it is checked out from the database, manipulations and changes made and rechecked in again. This procedure is a kind of QA to eliminate data errors. Here, the history stored in the database shows the changes and the person responsible. There are some control activities when checking the data back into the databases, in order to make sure that the geometry is correct. Other quality checks are conducted on buildings and buildings’ registry. So, QA is performed during any change or modification that is made on an object. Figure 6 shows the data production environment flow chart.

It is of high importance and help when producing reliable maps to know what has happened in the city and how it is evolving. The updating procedures at Stockholms Stad serve to fulfill this task. When changes are made in the city, the Base Map is the first one to be updated. The updates made in the Base Map, which are considered of interest for the City Map, are marked and stored in a so called Signal Map. These changes are for example a new building or a building that has been removed from its addressed location. When the marked objects have been used to update the City Map they are deleted from the Signal Map.

The same procedure is carried out when updating the 3D and County Map but now the updates are based on the changes in the City Map and not the Base Map. As before, the interesting updates/objects in the City Map are marked and stored in another Signal map, used for updating the 3D and County map and deleted after use. In the Signal Maps, there is a color coded labeling system. Green is used for new objects/buildings, red for deleted objects and
blue for modified ones. This updating process for the Base Map, City Map and 3D Map/County Map could be described as a chain where the links are: update data, use updated data and delete used updated data. The chain of the data updating process is illustrated in Figure 7.

![Data Updating Process Chain](image)

*Figure 6. Flow Chart for Data Production Environment (Hedman, K., personal communication, 13th May 2013).*

![Updating Chain](image)

*Figure 7. Data Updating Process Chain (Hedman, K., personal communication, 13th May 2013).*
5.3 Sjöfartsverket

5.3.1 Introduction

Sjöfartsverket or the Swedish Maritime Administration has the role of providing and developing secure, efficient and environmental safe shipping routes with service day and night. Their main client is shipping or navigation. Other interests of the organization are for example ports, municipalities, tourist boating and other transportation means, logistic companies, governments and other nations, the armed forces, the coast guard and vendors that transport their products by sea.

The basic services provided by Sjöfartsverket are (Sjöfartsverkets verksamhetsbrochyr, 2012):

- **Fairways**: Ensuring that shipping lanes are open and well maintained (managing dredging and widening of fairways).
- **Pilotage**: Some fairways require pilot to ensure safety. Sjöfartsverket’s pilots are about 200 around the country, all with high nautical skills and good local knowledge.
- **Icebreaking**
- **Sea traffic information**: VTS (Vessel Traffic Service) central stations monitor and communicate with vessels that travel through heavy traffic areas.
- **Vessel reporting**: The vessel reporting system, Safe Sea Net Sweden, is open around the clock for captains and other officers to report mandatory information.
- **Sea and air rescue**: Searching and rescuing people in need, locating missing aircrafts and vessels within Swedish territory.
- **Hydrographic information**: Seafloor surveying, measuring depths and locating foreign objects. Data is stored in databases for producing nautical charts, both paper and electronic.

5.3.2 Analysis of the current QA process in Sjöfartsverket

Being ISO 9001 certified Sjöfartsverket has identified the boundaries of the management system and produced documented procedures and process maps of its operations. Within each of these procedures and processes there is a quality check activity performed. When producing data, charts and other products, there are different stages of quality assurance; on board during the data collection, before accepting delivered data, during chart editing, proofreading before publishing. All of these stages are in place in order to ensure that quality requirements are being met. This shows that the geodata goes through a continuous QA process from the surveying phase till producing charts and maps. Other quality actions and tools that are used are checklists, handbooks, reporting system for errors and suggestions, and internal and external audits to ensure that the quality system continues to work.

Sjöfartsverket’s main role is to manage the invisible. The importance of data quality has been involved in the organizational culture since it was founded. It is the responsibility of the organization to map the seafloor in order to guarantee safe passages of ships, boats and other sea transportation vehicles. Therefore, it cannot afford to publish incorrect or inaccurate data for it might result in disasters and possible loss of human lives. For this reason and others,
Sjöfartsverket strives to involve everyone in the QA process of their data and ensure that every employee fulfills their responsibilities.

In Sjöfartsverket, it is the responsibility of everyone across the organization to perform data QA in her/his assigned area, from the surveying phase through producing the final geodata for publishing. Data quality is of high importance and is involved in the main management policy of the organization. The reliability of produced data, charts and other geodata for publishing, is a serious matter within the organization and therefore continuously monitored and evaluated.

Some automated QA and QC tools are used within Sjöfartsverket but their function is to check consistency within the data rather than the content. The automated tools are mostly used in the final stage of production when electronic nautical charts are to be published and delivered to customers. The consistency check that is performed in these charts is, for example, to check if locations of light houses are correct.

One of these tools is an indicating system called Automatic Identification System (AIS), which could also be classified as a labeling tool. This is a quality check system that detects errors in data by cross-checking them against the checklists of the survey and other geodata processes. Another indicating tool is the red light signal in the data collection process when using GPS devices. The red light indicates the weak signal of the GPS that may lead to inaccurate data records. Here, when there is a red signal, the collected data (GPS reading) is not further processed and the data is recollected when a sufficient GPS signal has been reached.

Although Sjöfartsverket has established a strong QA process, there is room for improvements. First of all, the criteria for the quality ambition level could be clarified. Some employees are invested in every detail of data quality while others are more liberal. New tools for automatic QA and QC could be implemented but not until the tools they have already are improved to their fullest potential. Finally, communication and feedback between different departments or sections, especially during the proofreading, could be enhanced. Doing so, it could be possible to minimize steps within each quality assurance phase and thus optimizing the procedure. Also, it could prevent mistakes from happening again and errors from reemerging.

Sjöfartsverket’s QA process is based on the S-series standards, which are developed by the International Hydrographic Organization (IHO). Those standards are in compliance with the ISO 9001 and other QA related series standards. For hydrographic surveys the S-44 standard is used in order to have a responsive QA process. As mentioned before, the organization is ISO 9001 certified, so the entire QA process is compliant with ISO standards. Now, the IHO is developing a package of S-series standards for the Hydrographic Organizations worldwide, called S-100.

The operations of Sjöfartsverket differ greatly from other Swedish land surveying organizations. In consequence, they prefer to compare themselves with other hydrographical offices and colleagues in other countries. The organization has performed benchmarking with hydrographical offices in neighboring Nordic countries; Denmark, Finland and Norway,
especially in regards to the surveying procedures where it is continuously sought to develop more efficient measurement methods. Since Sjöfartsverket collaborates mostly with international hydrographical offices the norm is about the same. Most hydrographical offices follow the same standards although the interpretation of them might differ between countries.

### 5.3.3 Process maps

A simple process of the general data flow in the organization is shown in Figure 8. It starts from field operations by conducting a survey process that serves as the data source for collecting data for processing. The survey output data is sent to the information management system for processing which results in storing the final quality controlled product. There are two main databases for storing and publishing data, the depth database and the chart database. The depth database contains detailed information about everything related to the geodata products, while the chart database is more specific and contains information that will be used when producing charts. The data in the chart database is regularly updated, and to keep the customers up to date, Sjöfartsverket informs them about those updates almost immediately, usually within a day.

![General Data Flow in Sjöfartsverket](image)

*Figure 8. General Data Flow in Sjöfartsverket (Lindgren, R., personal communication, 15th May 2013).*

### 5.4 iMMAP

Information Management and Mine Action Programs (iMMAP) is a humanitarian information management organization of Vietnam Veteran of America Foundation (VVAF) based in Washington DC. The organization is involved in guiding the international responders to crisis, such as post war countries and natural disasters. In their mission, iMMAP’s emphasis is to provide support services to national and international authorities/agencies for decision making. This is done through management of reliable, on-time and appropriate strategic and geographic information. iMMAP has direct involvement in (iMMAP, n.d.):

- Development of Database and Geographic Information System (GIS)
- Analysis of information and operational decision support
- Advisories and training in information management and GIS
iMMAP controls the data quality by conducting the QA and QC processes in a simple manner and according to a clear process map but based on a very well established and firm quality policy. It is noticeable in the Information Management System for Mine Action (IMSMA) processes, as shown in the process map in Figure 9, that the QA/QC and data screening for maintaining the data quality, is a continuous process throughout the entire information management system. This ensures the data quality, according to the mentioned quality management policy, and allows better tracking and controlling for the process management.

In the IMSMA process, the operations (OPS) start to collect data and feed it into the task dossier. The data goes through a screening and assessment process to decide if the data quality is accepted for data entry, otherwise it will be returned to the IP’s section. Then the data is submitted to OPS for treatment and correction. This is set up as loop number one. After the task dossier is created, the data goes through a QA/QC and monitoring processes so as to purify them by screening. This is classified as loop number two.

Then the data is forwarded either to the Mine Risk Education/Victim Assistance/Clearance (MRE/VA/CL) section via loop number three, or to Information Management (IM) section via loop number four. Here, data is either cross referenced for MRE/VA/CL and forwarded to the links (labeling erroneous data) for screening or corrected in the spatial form out of the IM section. If the data is spatial, it will go through another phase for correction of the TPs (Turning Points of the addressed area in polygon shape), if needed. Otherwise the data quality is approved and the data is sent to the IMSMA main database for archiving and further use by end users.

![iMMAP IMSMA Processes](image)

Figure 9. Process Map for iMMAP Information Management System (Kajtazi, B., personal communication, 9th May 2013).
5.5 Comparison

The aim of this section is to compare the QA processes in the geodata organizations of this thesis, so as to analyze them and identify the shortcomings. Lantmäteriet, Stockholms Stad and Sjöfartsverket all emphasize greatly on the quality of the data that they produce. In these organizations the quality assurance of the data is the responsibility of everyone. In Lantmäteriet and Stockholms Stad it is the role of the same person to collect, check, correct and update the data but in Sjöfartsverket the operations are a bit different. However, every employee is involved in the data quality assurance procedures in one way or another in all organizations. In regards to the organizational cultures it can be stated that data quality is considered as a high priority. The organizations are highly customer oriented and customer satisfaction is one of their most valued objectives. The importance of data quality is particular present at Sjöfartsverket since publishing inaccurate maritime data may involve great risks which could result in accidents, injuries or even life threat.

During the analysis of the interviews performed for this study it became apparent that the QA of geodata is not an independent process but integrated into the whole management system of the organizations. Each organization has a unique approach to the QA and how data quality requirements are fulfilled. Lantmäteriet receives large amount of data from municipalities. The section that receives data deliveries carries out quality checks to ensure the quality is at the accepted level. If not, the municipalities are contacted for correcting the identified errors. For Stockholms Stad most of the QA is carried out at site where the measurements are performed. In Sjöfartsverket the data QA is more comprehensive and more firmly defined. Sjöfartsverket performs the QA at different stages of the data management process; at measurement sites, before data entry and publishing to name a few.

Lantmäteriet and Sjöfartsverket work on a national level while Stockholms Stad focuses on activities within the city of Stockholm which rarely extend beyond the city boundaries. Hence, the extent and type of the data that the organizations handle and produce differs. However, there are things that these organizations have in common when it comes to data QA. For example, they all have regular organized meetings where quality issues are discussed, new measurement methods identified and improvement suggestions proposed.

Each of the interviewed organization has both strengths and weaknesses. Lantmäteriet has a good motivation and support system for the municipalities that they communicate with and provides feedback to appropriate parties regarding data errors and requirements that have not been fulfilled. Furthermore, the organization has recently published a process map for analyzing and managing quality failures that could be considered as an ‘Is’ map for the data QA process. On the downside their current error document system is only understood and used by a few.

Stockholms Stad works according to Lean philosophy that reflects the care they have for data quality. Very good updating procedures exist for the maps produced at Stockholms Stad where a so called signal map is used to keep track on changes made in the Base Map. Also, in the Base Map it is possible to obtain information such as data accuracy, measurement method,
and people responsible. Aside from the quality check done at site and data entry there are not many other QA procedures carried out before publishing.

Sjöfartsverket is ISO 9001 certified and has therefore established clear QA procedures. Along with the procedures there are other QA tools used such as audits and a reporting system for errors and suggestions. The reporting system is used and understood by most, if not all, employees and when something is reported there is someone assigned to respond. Lack of feedback and communication could be identified as a drawback for there are errors repeating each other which could easily be prevented.

Concerning the international organization iMMAP, they have a clear and simple ‘Is’ map of their QA process as well as a specific quality policy accompanying their overall organizational policy. It is clear from their QA process that it is a continuous process throughout the information management system.

Along with the strengths and weaknesses discussed here above for each organization, it is possible to summarize more general disconnections that should apply to most organization in the same field. The general disconnections identified during the interviews include lack of a specific quality policy, a separate QA process map, a responsible for the QA process and communication between employees.
6 Recommendations

The purpose of this chapter is to design and recommend an improved QA process for data management at the geodata organizations. The recommendations for QA improvements are based on theoretical research and practical studies conducted in the course of this thesis study. The theoretical research was carried out through studies from the Project Management and Operational Development Program courses and other external related sources. The practical studies were conducted through interviews and meetings with three geodata organizations, Lantmäteriet, Stockholms Stad and Sjöfartsverket, and communicating with another GIS related organization, iMMAP.

Theoretically, according to ISO standards the organization should have a well defined QA process that is integrated into the organization management system and mapped as a continuous support process throughout the information management system. Here, the QA process clearly contributes to the maintenance of the information management system and sustains the data quality and reliability. According to the international standards, in order to have a QA process that functions in a systematic way within the information management system, there should be a clear quality policy developed in compliance with the organizational strategy and required data quality.

Based on the meetings, interviews and communication with the mentioned geodata organizations, it was found that the QA is practically an essential and functioning process within the management system. In Lantmäteriet and iMMAP, there is a clear process map that shows the way of performing the QA through the information management system. In Stockholms Stad, the QA process is not mapped but integrated to the data management process and conducted by all the staff rather than allocated personnel in a specified QA section. Sjöfartsverket has specified QA process along the entire data management system; it is the role of everyone to conduct QA in their areas of responsibility.

The importance of the QA process is the focus of everybody in the geodata organizations. In all the mentioned data management systems the QA processes, which are performed in different ways, are functioning. But the point is as discussed in the disconnections in the previous chapter; there is no clear independent QA process and policy to be integrated as a main support process into the data management system. Hence, the following recommendations could be in assistance to improve the quality assurance in the geodata organizations:

- Develop a specific quality policy within the main organization’s policy and strategy.
- Establish a well structured QA process that is mapped clearly and emerged from the quality policy.
- Assign a QA process responsible; specifically skilled staff or a section within the organization, to control and manage the process. This QA process responsible will monitor, evaluate and directly contribute to the process improvement.
- Initiate a communication mechanism that involves the entire staff in providing feedback and disseminating information regarding the quality and work procedures.
improvements. This will prevent repeating mistakes and keep the new comers informed.

Having a clear QA process and a quality policy will ensure easier controlling and finding quality failures or defects in the process management. As a main support process to the data management system, the QA process may play a big role for the improvement and promotion of the data quality. Compared to the currently used QA process in the geodata organizations, when implementing a new process, there is an obvious change that should be managed by a organization management team. This change is to be reflected in the main policy as well as the quality policy.

The result of this study is a new recommended QA process, classified as the ‘Should Map’, which should, if applicable, replace the QA ‘Is Map’. It is a simple and clear QA process mapped throughout the data management system that shows the continuity of checking the data quality from the collecting phase through data processing and storing phases.

As shown in Figure 10, the QA process begins in the data collection phase by checking if collected data fulfills predetermined criteria. That is usually done through checklists and with handbooks at collection sites. When data is delivered it is screened and assessed before entered into the system. Before storing the data, quality failures are identified and managed either by direct correction or via loop to the screening step. After the data is stored in the system it goes through another filtering sub-process of data updating, if necessary. If new data is needed for the updating, requests are sent to receive, screen and assess step. The final quality check is performed when data is ready for publishing to review if the data quality is approved. If not, the stored data is sent back to identify the quality failure for rectifying. After the data is published, there might be changes made on the ground that should be updated in the data store. This is the versioning stage after storing the data in the database that will be performed regularly; the data may not have any quality failures, only needing updating.

In the QA Should Map, there is a continuous monitoring, evaluating and improving process that follows the life cycle of the entire QA process. This can clearly be seen in Figure 11 but also as a support process in Figure 10. This should involve the entire organization staff in addition to external experts for conducting auditing and evaluations. In this way, the QA process will be under constant surveillance and improvement to ensure the data quality and maintain reliability.
Figure 10. The Should Map for the data QA process.
Figure 11. Continuous QA Process for Data Management Systems.
7 Conclusions

Data quality assurance is of great importance for geodata organizations in order to have reliable and high quality data for publishing. Geodata products are used for planning and other infrastructure businesses that need accuracy in their data. Sometimes poor data quality can impact the safety of citizens and businesses, such as the ones that deal with water ways data. Thus it is vital to maintain data quality and keep it as high as possible. In this context, the QA of data could be identified as necessary for promoting the quality of the data.

Although there is a continuous QA process in the data management systems of the studied organizations, there is still room for improvement. Currently, the QA process is functioning but integrated into the data management system in the geodata organizations. A specific quality assurance policy within the organizations’ main policies is not available, which might be the cause of the lack of a clear QA process map, at least in some of the systems. This makes controlling of the QA process to be cumbersome, and tracking the defects within the process to be difficult. Furthermore, measuring the QA process time here is not easy, which is a hamper for doing valuable improvements of the process by identifying the critical path/s.

All persons involved agree about improving the data quality assurance process, with the aim of having it as a systematic support process for the data management system. In order to achieve that, it is recommended to develop a quality policy along with the organization’s main policy within their strategy, which reflects the importance of continuity of the QA process. Also, mapping the process will play an essential role in performing a smooth and controllable QA process, with opening the space for making effective improvements.

When applying the recommended QA process that is outlined earlier as ‘Should Map’ in the data management system, the communication mechanism within the process contributes to the involvement of all the associated staff/sections in a direct communication during the data processing. Consequently, the discussion of the data quality will be simple or as a routine among them during the data updating and quality check activities. This would create a proper ground for overall involvement of data management staff in upgrading the data quality and enhancing the work activities along with the QA process.

In conclusion, if the recommended QA process is successfully applied in the organizations studied in this thesis it could prove to be applicable for geodata organizations countrywide. It could be used for promoting the data quality and maintaining the effectiveness and efficiency of the data management systems and work practices. Certainly, applying the new QA process will be a change from the currently used one, and thus should be planned and managed by the organizations. However, before implementation, it should be assessed and evaluated for each organization to ensure that the change is beneficial.

In the future, further studies could involve interviewing more geodata organizations, national and international, in order to gain a more comprehensive perspective on geodata management and QA processes. Furthermore, if the above discussed recommendations are implemented a follow up could be carried out in order to evaluate the potential achievements in the data management system.
References


Information Facility, Copenhagen.

of Graduate School of Design at Harvard University:

March 2013 from College of Earth and Mineral Sciences, The Pennsylvania State
University: https://www.e-education.psu.edu/natureofgeoinfo/c5_p3.html.

International Conference on Information Quality (ICIQ-04), MIT Sloan School of
Management, Cambridge, MA, USA.

Environmental Protection Agency, Washington, DC.

5G*. United States Environmental Protection Agency, Washington, DC.

http://www.esri.com/what-is-gis/overview#top_five_panel.

http://www.esri.com/what-is-gis/overview#can_do_panel.

services/.

Sweden’s infrastructure for geodata website: http://www.geodata.se/en/Join/Participation-
agreement2/.

Retrieved 24th March 2013 from GeoVITe’s webpage:
https://geodata.ethz.ch/geovite/tutorials/L1IntroToGeodata/en/html/unit_u1Intro.html.


Intetics. (n.d.). *Quality Assurance*. Retrieved 3rd April 2013 from Intetics’ webpage:


Appendix A
The benefits of GIS

Organizations of all sizes and any industry can make use of GIS. Nowadays, there is a noticeable recognition of the strategic and economic values of GIS. The following categories can illustrate the benefits of GIS (Esri, n.d. a):

- **Savings in costs and higher efficiency**: wide use of GIS for optimizing schedules of maintenance and continuous fleet movements. Through reduction in use of fuel and staff time, enhanced customer service and more efficient scheduling, a saving of 10% to 30%, in cost, can be made.

- **Better decision making**: GIS is known as go-to technology for better decision making about locations, such as selection of real estate site, route or corridors, oil extraction sites and evacuation plans. Through using GIS, the right decision becomes clearer which is vital for the success of the organization.

- **Enhanced communication**: Maps and visualizations, in GIS, are great assistance for understanding situations and telling stories. They are good visualization tools to improve communication between different departments, organizations and the public.

- **Better recordkeeping**: Many organizations, governmental and authorized private companies, have the responsibility to maintain reliable records about the change and status of geography. Hence, GIS provides a strong framework for management of such records with full reporting tools and transaction.

- **Geographical management**: GIS is very important for understanding the current status and near future changes, i.e. what is going on and what will happen later, in a specific geographic space. After an understanding of the situation has been established, recommendations or advices for actions can be made. This is a good methodology for renovating the organization’s operations.

The GIS provides a new way to look at the world around us. With GIS it is possible to (Esri, n.d. a):

- **Map Where Things Are**: this enables us find places that have the features that we are looking for and to see patterns.

- **Find What's Inside**: this is to know what is inside a specific area. By using GIS, we can monitor what's happening and to take a specific action. For example, a district traffic police would monitor over-speed drives to find out if a case is within 1,000 feet of a school; if so, harder penalties apply.

- **Find What's Nearby**: helps us to find out what is happening within a set distance of a feature by mapping what is nearby.
• **Map Quantities:** we use map quantities to find places that meet our criteria and also to take necessary action. For example, public transport officials might want to map the numbers of busses per 1,000 people in each census tract to identify which areas are adequately served, and which are not.

• **Map Densities:** this helps in measuring the number of features using a uniform areal unit so as to see the distribution clearly. This is especially useful when mapping areas, such as census tracts or counties, which change greatly in size. For example, the number of people per census tract, the larger tracts might have more people than smaller ones but some smaller tracts might have more people per square kilometer; a higher density.

• **Map Change:** by mapping the change/s in an area we can anticipate future conditions, decide on necessary actions, or to evaluate the result of an action or policy. Thus, we can gain insight into how the changed things behave.

**Formats for Geographic Data**

Data can be described as references to observations and measurements related to real-world or imaginary entities and observable facts. Primary means for organizing spatial data are:

- Tables
- Vector feature classes
- Raster layers

These basic data provide a foreseeable means of organization that allows their tools to exchange information, engage information with operations and to discover associations regarding information from different data sources (Cote, n.d.).

**Essential factors about geodata quality**

Nowadays, geographic information plays a vital role in our modern society; especially when reliable and updated spatial information is needed, e.g. in disaster response systems. Also, geographic information is an essential input for decision making process of government’s entities such as urban and traffic planning, market exploration, location analysis or real estate analysis (Yang, n.d.)

Geodata production is much more than just generating a point, line or a surface object but should present a true situation about attributes of a spatial object. The quality of geodata can be articulated as follows (IABG, n.d. and Yang, n.d):
Table 3. Essential factors about geodata quality (Yang, n.d.).

<table>
<thead>
<tr>
<th>Factors</th>
<th>Quality considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location data:</td>
<td>• A group of data, such as (longitude, latitude, elevation) and (x, y, z).</td>
</tr>
<tr>
<td></td>
<td>• Relationship between/among Spatial objects, such as joint, orientation, etc.</td>
</tr>
<tr>
<td></td>
<td>• Directions; such as azimuth, lower/higher</td>
</tr>
<tr>
<td>Attribute data:</td>
<td>• Owner of a certain area</td>
</tr>
<tr>
<td></td>
<td>• Area and perimeter of a certain spatial object</td>
</tr>
<tr>
<td></td>
<td>• Color, vegetation type</td>
</tr>
<tr>
<td>Data Management:</td>
<td>• Structure of spatial database</td>
</tr>
<tr>
<td></td>
<td>• Data input, edit</td>
</tr>
<tr>
<td></td>
<td>• Queries according to objects or attributes</td>
</tr>
<tr>
<td>Data reproducing:</td>
<td>• Density of population/pollutions</td>
</tr>
<tr>
<td></td>
<td>• Add up/average of some attributes of a certain spatial object</td>
</tr>
<tr>
<td></td>
<td>• Processing model</td>
</tr>
<tr>
<td>Uncertainties:</td>
<td>• Fuzzy boundary</td>
</tr>
<tr>
<td></td>
<td>• Gradual changing boundary</td>
</tr>
<tr>
<td></td>
<td>• Determinant attribute according to different criteria, such as amount or area</td>
</tr>
<tr>
<td></td>
<td>• Round numbers</td>
</tr>
<tr>
<td>Effectiveness for real time:</td>
<td>• The existence of a certain object must be existed in a certain place (spatial attribute) and I a certain period/time (temporal attribute)</td>
</tr>
<tr>
<td>Data output:</td>
<td>• Figure</td>
</tr>
<tr>
<td></td>
<td>• Table</td>
</tr>
<tr>
<td></td>
<td>• 2D or 3D</td>
</tr>
<tr>
<td></td>
<td>• Visualization</td>
</tr>
<tr>
<td>Sharing exchange:</td>
<td>• Share</td>
</tr>
<tr>
<td></td>
<td>• Exchange</td>
</tr>
<tr>
<td></td>
<td>• Check</td>
</tr>
<tr>
<td>Legal dispute:</td>
<td>• Loss/lack aroused by using spatial data</td>
</tr>
<tr>
<td></td>
<td>• Legal and illegal</td>
</tr>
</tbody>
</table>

As a conclusion, the primary factors for evaluating the reliability of geodata could be classified as the location accuracy, the effectiveness for real time the processing model (Yang, n.d.).
Appendix B

Below are the questions used for the interviews at Lantmäteriet, Stockholms Stad and Sjöfartsverket. These questions are designed to shed some light on their current QA process in data management in order to analyze it and suggest some recommendations. The interview is thought as an open discussion about the QA process, where the questions are used as discussion topics. Some questions even lead to other questions that are not in the list but naturally emerge during the interview.

The interview: Questions and discussion points

- Have you ever mapped and analyzed the geodata quality assurance (QA) before? What is the IS process map for QA? How do you manage the QA?

  **Related topics:**
  - Is there specific diagram available for data quality assurance process?
  - Are different sections or departments of your organization working on assuring quality in the data? What are their different roles in assuring quality?

- Is this the most efficient and effective process, achieving the QA goal?

  **Related topics:**
  - Do you think that the quality assurance for data needs to be improved in your organization? If so, why do you think so? Do have any ideas or suggestions on how the quality assurance can be improved?

- Is the QA process involved in the organization’s culture?

  **Related topics:**
  - Is the importance of data quality being communicated across the organization, for example in regular/weekly meetings? Are you discussing errors detected in the data on regular basis?

- Are you using any QA management tools? Automated QC and QA? What software exists?

- What QA benchmarks/comparisons to other organizations have you created so far?

  **Related topics:**
  - Do you think you could benefit or improve your quality assurance procedures by comparing your work with other organizations and sharing information? If so, what kind of benefits do you think you could gain?

- Does the QA process design involve everyone in the organization?

- Do you have any quality indicators to detect errors or a flagging system that lets you know if there is an error in the data?
• Do you have any labeling of the data? For example to let users know that specific data is not accurate, needs cleaning, etc.

  Related topics:
  ▪ What kind of labeling system is there? Are there any special icons used or messages?

• Who is responsible for the QA process? Is everybody responsible for the QA of their data?

• What should be quality assured? What parts of the process?

• Are you using any standards (for example ISO) in you QA process today? Do you have any specific policies or strategies?

  Related topics:
  ▪ If so, what ISO standards have been implemented and you are using now?
  ▪ If so, what ISO standards do you plan to implement in the future?

• What is the general data flow in your organization?
Appendix C

Additional information about Lantmäteriet’s operation

Lantmäteriet’s main activity is forming and changing real property units through cadastral procedures. An example of such procedure is supporting landowners that hope to develop their properties in some ways. These developments could be among others related to roads, water resources or wastewater facilities (Lantmäteriet Annual Report, 2012:27). Lantmäteriet is accountable for the Real Property Register but that contains nationwide information about Sweden’s 3.2 million real property units. This information is very important to the banking system as they account on real properties as collateral for unpaid loans. Other clients of Lantmäteriet include the Swedish Environmental Protection Agency, the Swedish Armed Forces, the National Rail Administration, the Swedish Transport Administration, local government authorities and private persons. It should also be mentioned that Lantmäteriet’s activities, their geographic and real property data is essential for many services, especially for emergency and rescue services; police, fire fighting and ambulance services (Lantmäteriet, 2013).

Today, Lantmäteriet uses aerial photography and laser scanning to produce maps and models. Most maps are in two dimensions but with the development of technology three dimensional maps are being made by combining aerial photos and laser scanning data. Lantmäteriet has been using aerial photography for producing national maps since 1937. The organization has over one million analogue aerial photographs in their data storage but today all aerial photography is digital. Around 25,000-30,000 aerial photographs are taken each year, covering around 33 percent of Sweden. Historic aerial photography can be very useful for monitoring changes in the environment such as forest growth and water levels. These changes may be indicative of pollution or even climate change (Lantmäteriet Annual Report, 2012:8).

In recent times, Lantmäteriet has begun to use airborne laser scanning to build an elevation model of Sweden. Presently, around 70 percent of Sweden has already been laser scanned. The laser scanning is performed in 2,000 meter height where a laser signal is sent from a sensor in the airplane and the time taken for the signal to reflect of the earth and return to the airplane. The result is a bundle of location points, each assigned x-, y- and z- coordinates. This elevation model can prove its usefulness in many ways such as planning buildings and other infrastructure undertakings but its most important purpose is within risk management. That said, the model can be used to study the potential consequences of climate change for example large scale flooding or landslides due to excessive rainfall (Lantmäteriet Annual Report, 2012:12).

The Swedish government has placed great emphasis on national cooperation between organizations and recently assigned Lantmäteriet the task of coordinating operations within the geodata sector of Sweden. This responsibility involves establishing and overseeing a national geodata infrastructure in collaboration with local governmental authorities, county boards and other official organizations. The objective of this work is to enhance the access to geodata and interchange of information (Annual report, 2012:28). The Geodata Cooperation
Agreement is a product of this infrastructure development. The agreement serves as a channel through which authorities have access to digital geodata. This digital geodata is made available by different organizations through the Geodata Portal which presents available geodata and the responsible organizations for the data. Presently, there are 119 participants in the Geodata Cooperation Agreement that offer each other access to their geodata in exchange for annual fee (Geodata 2011, Geodata 2012 and Lantmäteriet Annual Report, 2012:28).

As for international assignments, Lantmäteriet contributes to social development in several countries where they aid in building well functioning real property administration systems to keep track on ownerships. Lantmäteriet recognizes that ownership is fundamental for all society development infrastructures and during the year 2012 it was engaged in projects in Albania, Azerbaijan, Bosnia, Botswana, Georgia, Jordan, Kenya, China, Namibia, Serbia and Belaru (Lantmäteriet, 2013).
Appendix D

Additional process maps for Lantmäteriet

Main process map for information management

Figure 12 illustrates the main process map for the information management system in Lantmäteriet. The first main process, control process, flows through five different sub-processes; starting by data capture and analysis, through strategic planning, planning operations, controlling operations and ending by follow up and improvement of the entire information management system.

The other main process, in the information management system, is driving/running process. Here, the idea is to develop the information management so that it functions smoothly in order to manage the sub-processes in a flowing way. In this process, the data quality should be defined clearly; data processing and management will result in the final product (geodata) for publishing. The running or driving process in this information management system is composed of three sub-processes; collection and insertion of information, information storing and finally product delivery.

Collecting and inserting information is a sub-process in the driving process. When the production plan is developed, input is supplied through a signal from the sender and a task order is issued to collect information. Collecting information involves taking pictures, measuring heights, GGD & buildings, documenting locations names and mapping the data using geodesy methodologies. This sub-process is followed by QA procedures so as to have totally or partially controlled geodata ready for the storing process. After inserting information and conducting quality assurance, another sub-process follows; it is storing geodata, where information and data is processed, stored and archived. This sub-process involves managing geodetic reference system (geodetic infrastructure), executing tasks from authorities, such as location name decision and database permit, and coordinating geodata, which is considered as infrastructure for geodata. The last sub-process, in driving process, is delivery of analog and digital product and services to be available for customers and users. Managing orders from the customers will be done prior to this stage, in order to enable delivery of the product accordingly.

For every functioning system, in addition to the main processes that are involved, there should be support process/es that complement the main processes so as to sustain the system and enable efficiency and effectiveness. Examples of those support processes are maintenance and financing of specific operations. The support process, here in the information management system, involves some sub-processes. Starting with performing system management and managing quality defects, legal issues, relations and support activities.
**Process for updating ABT**

Figure 13 illustrates the process for receiving, verifying and updating data about addresses, buildings and topography (ABT). The process starts with an agreement between the municipality and Lantmäteriet containing what data is expected of the municipality, how it should be presented and when it should be handed in. Lantmäteriet either receives data from the municipalities according to the agreement or requests for it as needed. When a data delivery has been received and accepted a visual inspection and later a more thorough examination is performed where the content, structure, attributes, overlaps and duplicates are studied. If errors in the data are found at this stage, the responsible municipality is contacted and a new delivery or a correction is requested. This is followed by an updating procedure which differs regarding the type of data is being considered, building geometry or other topography. Several decisions points are placed throughout the updating process where merging errors, conflicts and collisions are analyzed and appropriate actions taken. Once the updating procedure is completed the data is sent to relevant parties.

**Support process for BAL**

Figure 14 illustrates the support process for buildings, addresses and apartments (BAL). The objective is to provide support and advices to the municipalities within this area of operations. When a new case of addressing and registration comes in then the support sub-process will start, through a continuous process, by a provision officer to prepare it for further processing. The case will be assessed by the officer so as to be ready for investigation and correction, if needed. This is the quality check that the involved people will do when receiving the case. The case might go through another process for investigation and completion, in both receiving and preparation phases, and sometimes needs expert assistance for correction and completion. Hereafter, the case will either be sent back for further investigation or feedback or will be corrected during the quality check phase, then it will be reconnected for feedback and updates before case management completed. So, the final stage of case management completion is going through a continuous process in analysis, quality assurance, and investigation & feedback phases.
Figure 12. Main process map for information management at Lantmäteriet (Paasch, J., personal communication, 16th April 2013).
Figure 13. The Process for receiving, verifying and updating data about ABT (Paasch, J., personal communication, 16th April 2013).
Figure 14. The support process for buildings, addresses and apartments (BAL) (Paasch, J., personal communication, 16th April 2013).