Usability of Networked Information

What influences the usability and how it can be measured

DISSERTATION

Wouter J. Könst

submitted in part fulfilment of the requirements for the degree of
Master of Science in Geographical Information Systems (UNIGIS)
Faculty of Earth and Life Sciences
Vrije Universiteit Amsterdam
The Netherlands
October 2013

[This page is intentionally left blank]

Abstract

Wouter J. Könst, 2013

Usability of Networked Information

What influences the usability and how it can be measured

Networked Information (NI) is generated real-time from at least one source available via the internet and it is distributed online. It is widely used, privately, commercially and by the government and affects different areas of society like social structures, business and public safety.

Therefor the usability of NI, the degree to which the product meets the requirements of the user, should be defined and made measurable. This provides the user with means to select the product that best fits its needs and enables NI producers to improve the usability of their products.

From this background, the following research question was formulated: Which aspects of networked information influence its usability and can the aspects be used to measure and compare the usability of networked information products?

Based on literature study, five aspects are identified, namely privacy, quality, reliability, security and transparency. The primary use of NI is to assist in decision making. And therefor it has to be legal and of good quality and the service needs to be secure and reliable. Also, to be able to assess the above, the producer needs to be willing to provide the user with inside information.

Information on which the level of usability of networked information is based, is obtained by means of an interview. Because NI are used by a wide range of customers, also those without knowledge about services and or geo-information, questions in the interview are aimed at company processes.

In this thesis, the selected aspects and the method of obtaining insight in the usability of networked information are discussed.

[This page is intentionally left blank]

Table of contents

Lists of figures

List of graphs

List of tables

Disclaimer

The results presented in this thesis are based on my own research at the Faculty of Earth and Life Sciences of the Vrije Universiteit Amsterdam.

All assistance received from other individuals and organizations has been acknowledged and full reference is made to all published and unpublished sources.

This thesis has not been submitted previously for a degree at any institution.

Signed:

Den Haag, 30 October 2013,

W.J. (Wouter) Könst

[This page is intentionally left blank]

Acknowledgements

Firstly I want to thank the people who helped me in the various stages of my research. Jasper Dekkers of the VU and Cees Guikers of Bridgis, you provided me with a topic for my thesis, and looking back, it has proven to be a very interesting subject. Niels van Manen of the VU, although we only met recently, thank you for the nice conversation and informing me about what to expect at graduation. Erik van der Zee of Geonovum, you helped me at the start of my thesis, by introducing me to the world of services. Your information helped me get a grip on the topic and narrow my research.

Special thanks to my supervisor, Rob van de Velde. You were very supportive throughout the process and shared your knowledge and experience with me. I liked sparring with you and your critiques have helped me improve my work.

Furthermore I would like to thank all participants of my interview, for their time and openness about their organization and their products and my employer, the provincie Zuid-Holland, for providing me with facilities to start and complete this study.

Thank you family and friends. Mom and dad, I did it ☺. Anne and Nadja, Paul and Adeline, Erik, Caroline and Robert, thanks for your interest and encouragements in the past years.
I want to thank and remember my mother in law, José. Your strength and perseverance are an example to me.

Most of all, I want to thank my family, Danielle, Karlijn and Maartje. In the past few years you had to do without partner and father for quite some evenings and weekends. Thank you for supporting me and giving me the space and time to work on my study. Love you.

CHAPTER 1
Introduction

1.1 Introduction to Networked Information

… “the move to a communications environment built on cheap processors with high computation capabilities, interconnected in a pervasive network - the phenomenon we associate with the Internet” (Benkler, 2006, p.3), has made the outsourcing of ICT possible through service oriented architectures (SOA). Software, platform and infrastructure as a service (respectively SaaS, PaaS and IaaS) are accepted, but still many organizations kept their information management internal. Organizations had to spend money and resources to capture, process, save and off course maintain data, so it could be used in the business.

This is changing. Although some organizations choose not to change their data infrastructure because of strategical (e.g. the data is too valuable, like with the Pentagon), organizational (e.g. change causes many layoffs) or technical (e.g. lack of equipment) reasons, an increasing number of organizations is putting all or part of their information management on the market, which has led to a new member of the ‘as a service’ family: Data as a Service (DaaS). DaaS “… involves three main areas of interest: providing commercially available data through some form of interface, wrapping up a firm’s own data through a service layer, and data centre infrastructure services offered as a product. Some DaaS systems even offer a mixture of the three” (Williams, 2012).

In the past decade the amount of data available through the internet has strongly increased. Governments are sharing more and more data they have available and at the same time individuals produce an enormous amount of social data.

Longley, Goodchild, Maguire and Rhind (2005, p.12) mention that a “… characteristic of information is that it is easy to add value to it through processing, and through merger with other information”. And that is exactly what companies are doing. They use the available (online) data to create a product that they can sell and make profit of: information.
By combining and processing the available online data with or without data from other sources, companies can meet a market demand and make the produced information available as a service. The product is Networked Information (NI). It is generated real-time from at least one source available via the internet and it is distributed online. It is a result of a Model as a Service (MaaS), which consists of the elements data, service and application. Data is supplied to the service by Open Geospatial Consortium (OGC) standards, Web Feature Service (WFS) for vector data, Web Coverage Service (WCS) for raster data and a Sensor Observation Service (SOS) for sensor data. Dwivedi and Kulkarni (2008, p.615) describe the process of creating NI, which “entails capturing raw information across sources, its storage and validation and providing analytical reports to the clients” (figure 1). The

process has three actors. The supplier of the data used to generate the NI, the producer of the NI and the client, who buys and uses the NI. The communication, including checks on for instance quality and privacy, is mainly (if not entirely) done by machines. This creates a large dependency on hard- and software.

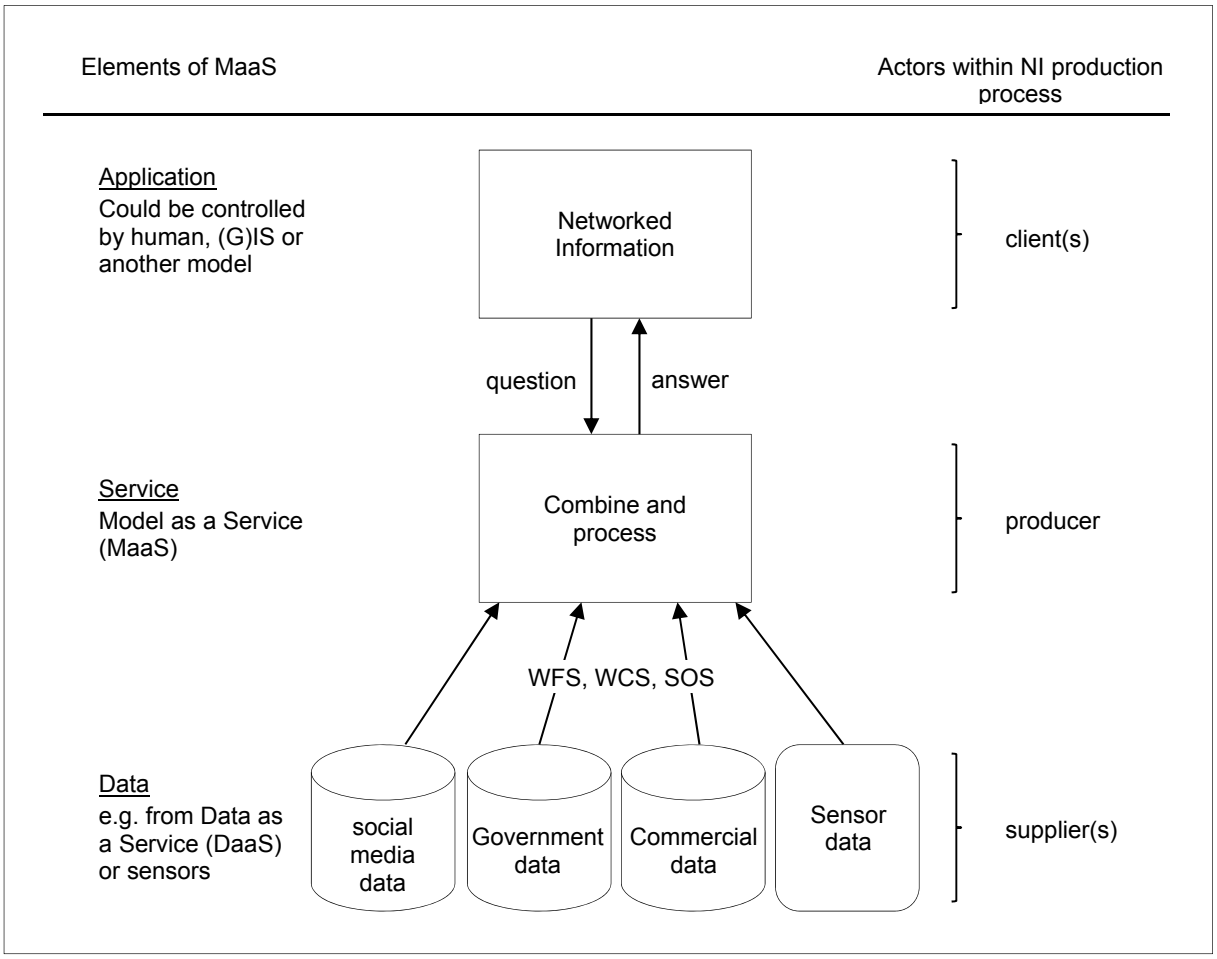


Figure 1: Schematic structure of Model as a Service (MaaS) and its relation to NI. The elements of Maas are shown on the left and the involved actors of the NI production process are shown on the right.

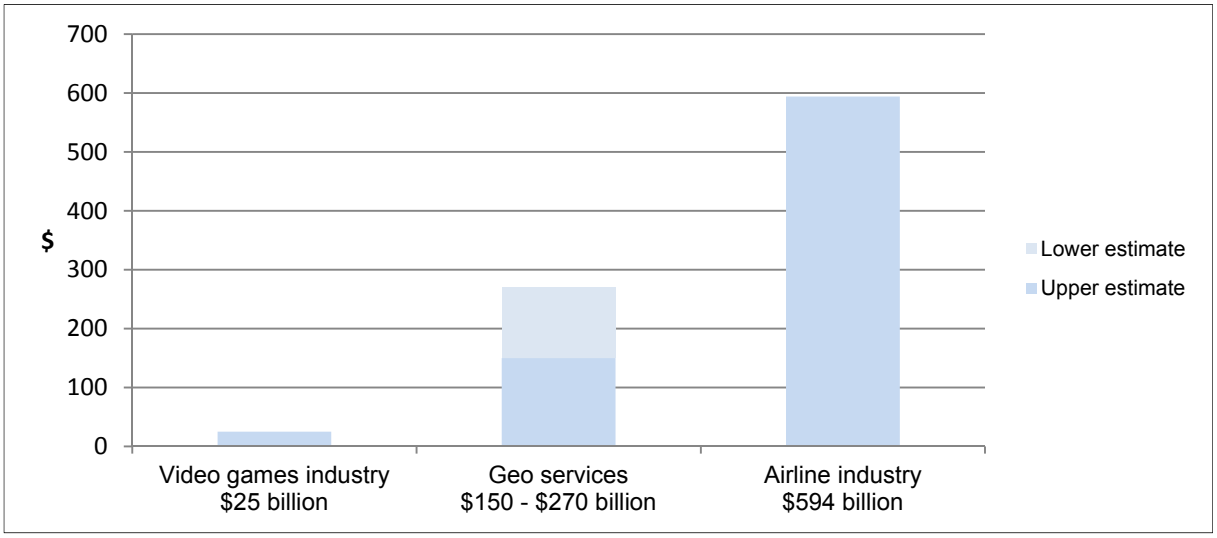
NI resembles DaaS (with an emphasis on the D for data) but NI is an end product, it is ready to use for a specific purpose. Longley, et al. (2005, p.12) describe two ways in which the term information can be used. First, it "… can be treated as devoid of meaning, and therefore as essentially synonymous with data" and second that it can be described as "…data serving some purpose, or data that have been given some degree of interpretation". In this thesis, from now on the latter description will be used in which data is the raw material information is made of.

The benefits of NI are much like the benefits for DaaS. First, it is cost effective, you buy what you need. Enterprises can "… focus on what they do well - run their business. Money does not need to be misspent building an internal Data Management team, if that team is not an important requirement of doing business" (Williams, 2012). Second, it increases flexibility because accessing the data is location independent. And third, "access to the data is controlled through the data services, which tends to improve data quality because there is a single point for updates. Once those services are

tested thoroughly, they only need to be regression tested if they remain unchanged for the next deployment" (Rajesh, Swapna and Shylender Reddy, 2012, p.28).

According to Rajesh, Swapna and Shylender Reddy (2012, p.28) the drawbacks are that the client is dependent "on the service provider's ability to avoid server downtime" and that "the consumer is really just "renting" the data, … generally the data is not available for download".

Networked Information is widely used, privately, commercially and by the government. In 2012, Oxera was commissioned by Google to estimate the worldwide economic impact of Geo services. They came to an amount in the range of $150 to $270 billion. Although Oxera indicates that the estimate is based on combining information from different sources and "should be considered indicative of the likely magnitude of impacts" (Oxera, 2012, p.iii), these are still huge figures, even compared to other industries (graph 1)



Graph 1: Estimates of global revenues of Geo services compared to other industries (from: Oxera, 2012, p.i).

Some products are used daily by large amounts of people. These products have a great social impact just because of the amount of users e.g. buienradar.nl or filewissel.nl. Other products have a higher safety or economic impact because of use in healthcare or disaster management (saving lives) or because companies are depending on the information for their business (making money). This means information should be usable for a specific purpose (e.g. making money), which involves "…making useful data available to key users in a timely, secure, and cost effective manner … the goal is to enable broad and timely self-service access to business critical-information" (Delphix, 2011, p.3). In all cases high demands are made on the different aspects of the Networked Information product.

1.2 Aims and objectives

This thesis is a study to the usability of Networked Information and how this might be measured. Regarding to research, the Unigis MSc guidelines refers to The Oxford English Dictionary (Oxford University Press, 2012) which states research and development are "work directed towards the innovation, introduction and improvement of products and processes". This study is about improvement of the understanding of the quality and reliability of NI and thus the deployment of the product.
The phenomenon NI is emerging and its use is growing rapidly, privately as well as commercially and in government (e.g. risk management). This means that the social, economic and safety dependency on NI products is growing. That's why it is important to know what makes the information fit for use.
Ultimately, if the usability of the NI product can be measured, and the aspects that influence it are known, than producers of NI will be motivated to improve the usability of their product.

Because NI is a service and it is about information use, the aspects involved are related to (geo) information in general and to service oriented architecture (SOA), specifically to data, information and model as a service for the distributed and service parts.

Main question:
Which aspects of networked information influence its usability and can the aspects be used to measure and compare the usability of networked information products?

Sub-questions:
- What is usability in relation to NI?
- What requirements does NI need to have to be fit for use?
- How can the requirements for the fitness-for-purpose of NI be measured?
- To what extent can the producer of NI positively influence these aspects?

1.3 Research limitations and delimitations

This thesis focusses on NI products that have a geographical component. This has two reasons. One is that the use of information with a geographical component is so common that most people do not even realize that they are. It is taken for granted that location related information is accessible anywhere at any time. And it is. "Almost all human activities and decisions involve a geographic component" (Longley, et al., 2005, p.8). This is true for individuals and also for companies and governments. The second reason for focussing on NI with a geographical component is that "the UNIGIS Programme offers ample opportunities for professionals to acquire knowledge and skills in the field of GIS" (Unigis education, 2012).

The review will quantify the performance gaps between different products. And because not only specialized organizations use NI, everyone has to be able to assess the usability of the information in relation to a specific purpose, without specialized hardware, software, staff or knowledge. For this, the focus is on gathering information about the NI itself and the company's internal processes and not on extensive data checks.
Therefore, the information that the NI producers deliver is very important. It is assumed that the participants (producers of NI) will truthfully answer the survey questions. But there may well be a limitation regarding business sensitive information (e.g. server downtime). Companies might not be inclined to share this kind of information and another source of information has to be found (e.g. theoretical guaranteed values specified in the SLA).

The sample used in this study is randomly selected from the population, with the exception that no two selected products serve the same purpose. The hope is that participants (being the producers of the NI) are more inclined to provide information when their direct competitors do not participate in this research.

Usability benchmarking is a well-known tool that is used by companies to check if their products are good enough and meet consumer demands better than their competitors' products do. Different companies and organizations offer competitive usability benchmarking services (e.g. Userfocus, TecED and the Design and Usability Center at Bentley University). Benchmarking is a tool for evaluation and improvement, with which products are compared to other products which are "recognised as the best within the area" (Andersen and Pettersen, 1996, p.3). But, because the method of benchmarking in this study is used purely as a manner to implement the theory of measuring the usability of a NI product according to the described aspects, instead of looking for and comparing to the best product, the NI products will be compared to the baseline that is defined as a result of the literature study.

The subject is the usability of the networked information (e.g. quality, legal issues and reliability) and not the software, and therefore user experience (e.g. layout or satisfaction of user requirements) is not part of the research. When possible, users will however be contacted to compare their own experiences with the NI product to the information given by the producers.

1.4 Research methodology

First relevant literature is studied to define 'usability' in relation to NI and to find the aspects of NI that influence the usability of a NI product. This is a theoretical approach to 'usability' of NI instead of doing (qualitative) research to what people understand it to be.

To implement the theory of measuring the usability of a NI product according to the described aspects, a product usability benchmarking is performed with seven products. Data about the NI products to be compared is accumulated through a combination of web research and a questionnaire.

The different NI products that are used in the benchmark are described in chapter 4 (who is the vendor, what is its use, who is using it, etc.).

1.5 Thesis outline

In chapter 1 NI is introduced and the design of the research is briefly described. In chapter 2, usability of NI is defined, together with the aspects that influence the usability. In chapter 3, the method used for the research is described and substantiated. The results of the research are included in chapter 4, followed by the conclusions, based on the results of the research and the literature review, and discussion in chapter 5. The graphical outline of this thesis is shown in figure 2.

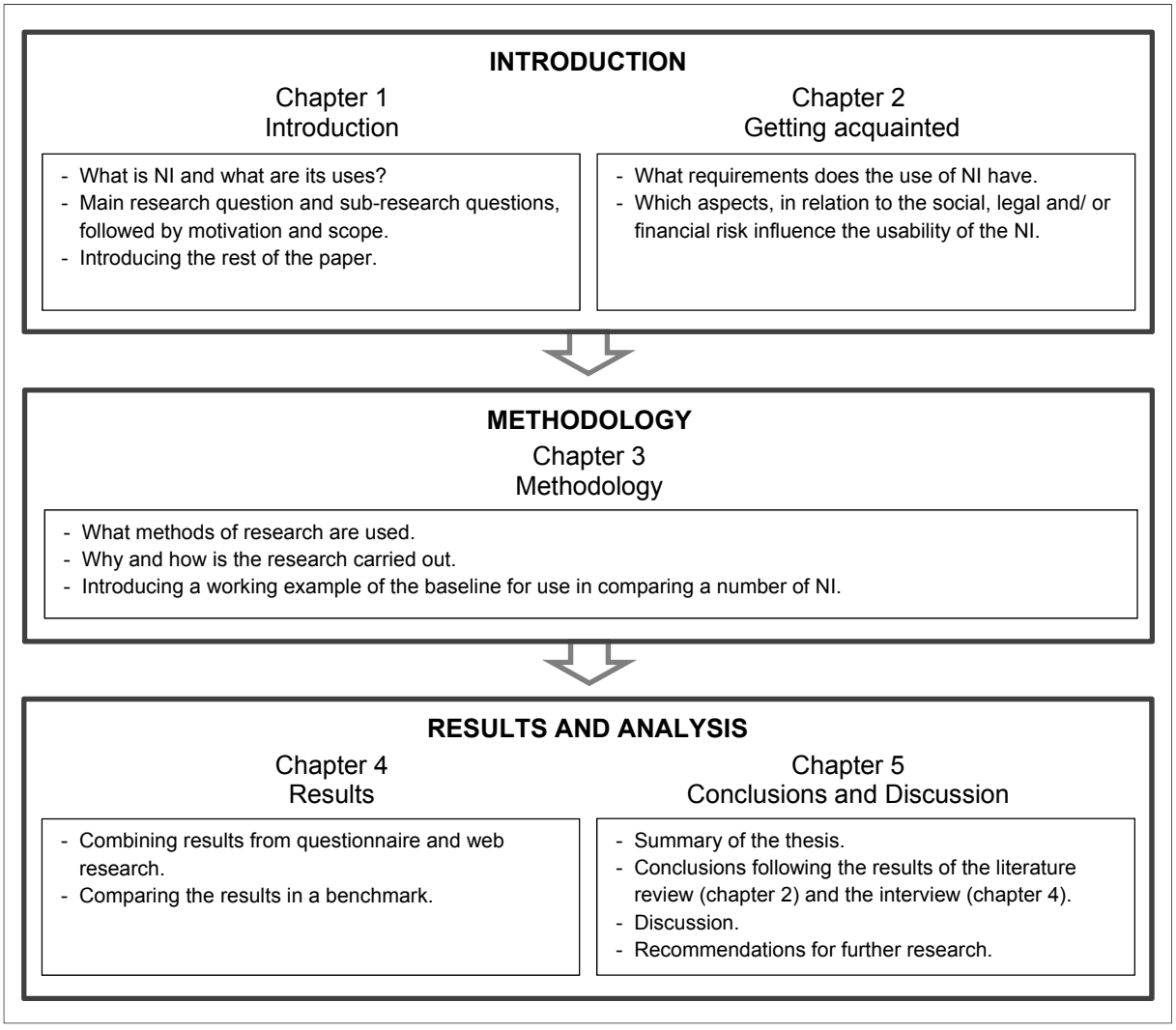


Figure 2: Thesis outline.

CHAPTER 2
What defines the usability of Networked information

2.1 Usability of networked information

“Usability is the degree to which something - software, hardware or anything else - is easy to use and a good fit for the people who use it” (User Experience Professionals' Association, 2012). Designing a usable product starts with usability engineering, which, according to User Experience Professionals' Association (2012) is “the disciplined application of usability practices to assess the needs and abilities of users, in conjunction with the business requirements, practices, and processes of an organization”. In other words, defining the usability of a product starts with finding out what the users’ needs are.

NI is **generated real-time** from **at least one source available via the internet** and it is **distributed online** and, in this dissertation, NI is a service which provides geographical information. Because of this, it has the combined characteristics of Geographical Information (GI) in general and service oriented architectures (SOA), specific data and information as a service. So identifying aspects that influence the usability of NI products requires looking at the needs of the users of Service Oriented SOA and GI. The primary use of NI is to assist in decision making. And for this, NI has to be legal and of good quality and the service needs to be secure and reliable. On top of that, to be able to assess the above, the producer needs to be willing to provide the user with inside information.

Because of the great quantity of digital available personal related data from social media and government and companies like Google and Microsoft who collect and use personal data in their business, privacy is a big issue in NI. Also data needs to be true, complete and up to date to be used in decision making. These three elements come together under the aspect quality of a data set.
Delphix (2011, p.3) mentions both timely and secure as required characteristics of DaaS. Clients should be able to access their data when they need it. In this context, timely is part of reliability, in that being able to access the data when needed means that the data should always be accessible. “The highest quality information is useless if it cannot be seen so that it can be acted upon” (Tortorella and Driscoll, 2005, p.3). The continuity of the data supply in case bankruptcy of the producer or a data supplier and availability of information by avoiding server downtime (Rajesh, Swapna and Shylender Reddy, 2012, p.28) are also part of reliability. Because information is being produced and kept external, by an outside party, clients must be able to rely on their sensitive data (e.g. business-critical information) being secure and unavailable for any unauthorized access and use.
For being able to assess the above aspects, the producer needs to display a certain amount of transparency. If clients are not given access to information relevant for measurement, they cannot assess the usability of the product.
With this, five aspects have been identified that influence the usability of NI; **P**rivacy, **Q**uality, **R**eliability, **S**ecurity and **T**ransparency.

2.2 Aspects that influence networked information usability

2.2.1 Privacy

“Although in principle it is possible to use GIS for any purpose, in practice it is often used for purposes that may be ethically questionable or may invade individual privacy... The technology may appear neutral, but it is always used in a social context” (Longley, et al., 2005, p.31). This is quite a statement but Longley has a point. The June 2004 edition of Reason Magazine had a personalized cover for each of the 40.000 subscribers. It showed a satellite photo of their own neighbourhood with their own house circled red (figure 3).

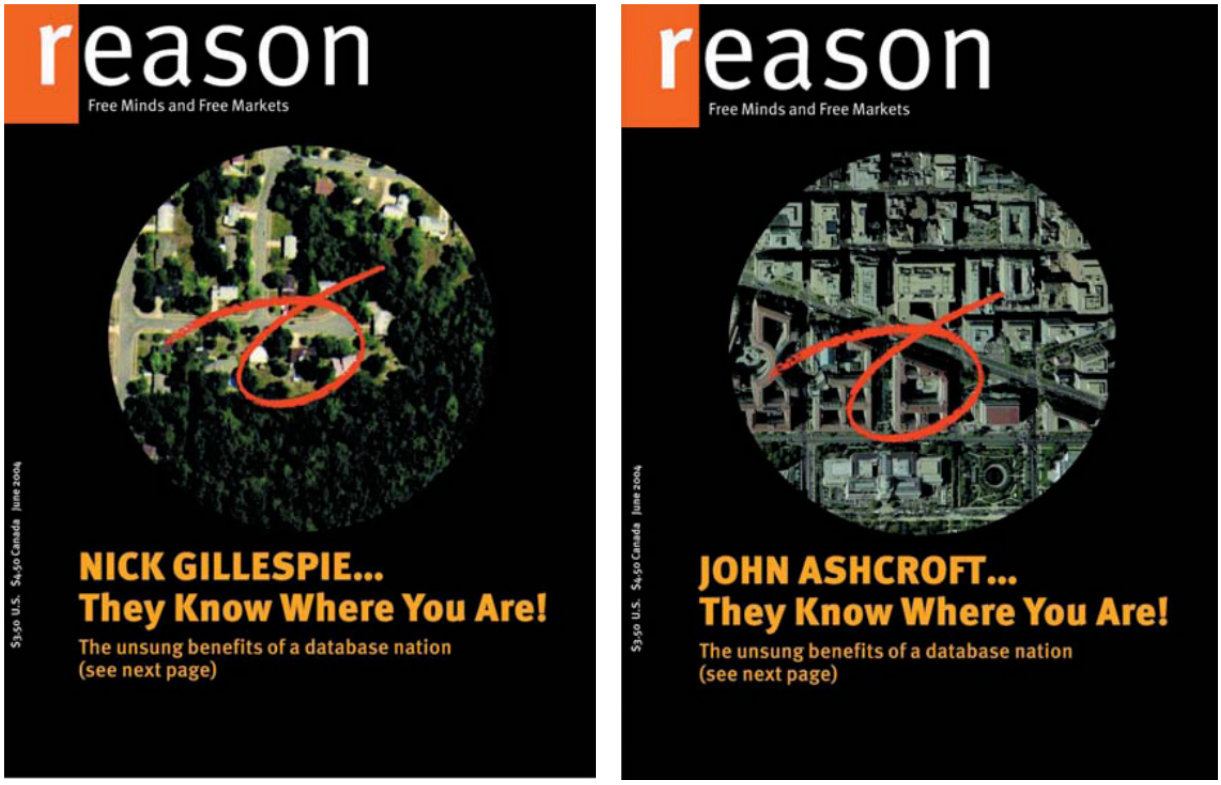


Figure 3: All 40.000 subscribers of Reason Magazine received a personalized June 2004 edition. It showed a satellite photo of their own neighbourhood with their own house circled red (from: Longley et al., 2005, p.421).

The point Reason Magazine wanted to make with the accompanying article “database nation” is the power and availability of data from databases (McCullagh, 2004). Nick Gillespie, editor in chief of Reason, said that "everybody, including our magazine, has been harping on the erosion of privacy and the fears of a database nation. It is a totally legit fear. But they make our lives unbelievably easier as well” (Carr, 2004).

Westin (1968 cited in Kagal, 2010, p.1) describes privacy as “the ability for people to determine for themselves “when, how, and to what extent, information about them is communicated to others””. This definition focuses on the access of information, which is outdated. Nowadays people choose to share their information (e.g. on social networks) or are obliged to do so (e.g. in healthcare). Thus it is less about the access of information, but rather what is done with the accessible information. Information is

money and sites like Facebook and LinkedIn use this development. They “exploit what economists call “network effects”. Because users are contributing information as well as utilizing information others supply, these sites become more valuable the more people are using them” (Abelson, Ledeen and Lewis, 2008, p.110).
So, after Warren and Brandeis (1890) *(privacy is “the right to be let alone”)* and the United Nations General Assembly (1948, article 12) *(“No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation”)*, Abelson (2010, p.17) suggests that in the (digital) information age privacy is no longer about private information that is communicated to others, but about private information that is used by others in ways that affect the individual concerning.

A way to depersonalize data is to aggregate it, for instance by displaying personal data by postal code area. Unfortunately, this method appears not to be as safe as assumed. As part of his study, Koot (2012, p.35) assessed the re-identifiability of entries in two Dutch research databases. The results showed that two thirds of the Dutch citizens can be unambiguously identified by combining date of birth and the four digit postal code.

The misuse of personal data has moral and legal implications for both the manufacturer and the user of NI. In many countries, (governmental) organizations for the protection of personal data have a responsibility to ensure that personal data are used carefully and secure. Also, Personal Data Protection Acts in different countries describe what is and what is not allowed concerning personal data. In whatever way the data is used, it has to be done in such a way that it is compliant with the Personal Data Protection Act. This ensures that both the manufacturer and the user of NI avoid being fined or risk reputation damage.

2.2.2 Quality

“The extent to which errors and other shortcomings of a data set affect decision making depends on the purpose for which the data is to be used. … Quality is often defined as ‘fitness for use’” (De By, 2001, p.401). This corresponds to the view of the Australian/New Zealand Standard (1994, p.v) that quality is the “totality of characteristics of a product that bear on its ability to satisfy stated and implied needs”. So data quality is a way to express the value of data for a particular task. Low quality data has a negative effect on decision making. This could have relative small consequences, for example when outdated route information in a navigation system is causing the user to drive in a wrong direction. It could have more serious consequences for instance, when a flood prediction model is using incorrect land height data or wind speed and direction, it could have an effect on evacuation plans.

So being able to assess the quality of a dataset is important for defining the usability of NI. De By (2001, p.401), Van Oort (2005, p.14) and Van Poperingen (2004, p.34) have together compared seven sources (table 1) in search of elements of data quality. In total they found 12 elements (table 2).

Source	R. de By (2001)	P. van Oort (2005)	S. van Poperingen (2004)
International Cartographic association's (ICA)	√	√	
S. Aronoff		√	
USA-SDTS		√	√
CEN/TC287		√	√
ISO/TC211		√	
Digital Geographic Information Exchange Standard (DIGEST)			√
Referentiemodel kwaliteit van geo-informatie (RAVI)			√

Table 1: Sources that each writer compared in search of elements of data quality.

Elements	R. de By (2001)	P. van Oort (2005)	S. van Poperingen (2004)
Positional accuracy Accuracy of coordinate values	√	√	√
Attribute accuracy "Accuracy of all attributes other than the positional and temporal attributes of a spatial data set" (Van Oort, 2005, p.15)	√	√	√
Temporal accuracy Time related information about the data set	√	√	√
Completeness A measure of the absence of data and the presence of excess data (Van Oort, 2005, p.15)	√	√	√
Logical consistency An indication of the graphic quality and topological integrity of a digital map (Federal Systems Management and Integration Center, 1996, p.14)	√	√	√
Lineage "Information on sources, update activity with dates, and processing steps that have transformed the data" (Federal Systems Management and Integration Center, 1996, p.14)	√	√	√
Meta-quality The quality of the quality definition (Aalders, 1998, p.7)		√	√
Homogeneity Description of expected or measured uniformity of quality parameters (Van Poperingen, 2004, p.25)		√	√
Semantic accuracy Meaning of and relationships between data		√	
Resolution The detail at which data are presented (Van Oort, 2005, p.12)		√	
Usage, purpose and constraints Helps assess fitness-for-use (Van Oort, 2005, p.16)		√	
Administration of users Helps assess fitness-for-use			√

Table 2: List of the elements of data quality that were found by De By, Van Oort and Van Poperingen.

In this thesis, quality is characterised by **correctness**, consisting of 'positional accuracy' and 'attribute accuracy', **completeness** and **temporal accuracy**. These are the elements that can be found in all the individual sources. The other elements are, as Van Poperingen (2004, p.35) puts it, "other quality aspects". They give information about the quality of the quality definition of a dataset ('meta-quality', 'homogeneity' and 'semantic accuracy'), or the properties or background of the dataset ('resolution', 'usage, purpose and constraints', 'Administration of users', 'logical consistency' and 'lineage'). As such, some of the "other quality aspects" can help assess the three quality characteristics.
Also, to help check if the content and metadata of datasets comply to defined standards, different tools have been developed by (standardizing) organizations (e.g. the Validator by Geonovum, the Web Testing Facility by Open Geospatial Consortium (OGC) and a schema validation by the U.S. Geoscience Information Network).

Correctness

Correctness indicates the degree to which something is good or true as does error indicate that something is wrong. This applies to both positional and attribute data. De By (2001, p.402) distinguishes "between errors in the source data and processing errors resulting from spatial analysis and modelling operations carried out by the system on the base data". Although the origin is different, the result is the same, i.e. NI of limited quality and thus limited usability. To get an indication of what has happened with the data, lineage can be of use. Van Oort (2005, p.14) describes lineage as "a description of the source material from which the data were derived, and the methods of derivation, including all transformations involved in the production process".
Within positional accuracy a distinction is made between relative positional accuracy, which "is the accuracy relative to other data in the same test data set" and absolute positional accuracy, which "is the accuracy of test coordinate values relative to matching reference coordinate values on the same coordinate reference system" (Van Oort, 2005 p.15). For most uses relative positional accuracy is sufficient when only one data layer is used, but "the fundamental tool of polygon overlay relies upon absolute positioning between layers" (Chrisman 1991, p.169). So when using more than one layer relative positional accuracy within one layer creates a risk of relative positional inaccuracy between layers.

Completeness

Mentioning completeness in relation to a data set, is generally considered with incompleteness. De By (2001, p.409) describes two 'flavours'. First, the lack of data as a result of a mistake during measurement (missing an object) and second missing attribute data, e.g. because of lack of knowledge. On the other hand, data shouldn't be 'over complete' by which the existence of doubles in a dataset is meant.
Whether a data set is complete depends on the agreed data model, which, obviously, is dependent on the intended use. A navigation program like TomTom does not need to give information about areas with flood risk or pipelines that transport e.g. gas or oil. On the other hand this is vital information for a risk map which goal is to inform people about the risks in their own vicinity (Risicokaart, 2013).

Completeness of groundcover (as is demanded with the Dutch registration large scale topography (Basisregistratie Grootschalige Topografie, BGT) is easy to check, holes in the surface topography are simply not allowed. Organising objects (e.g. trees and lampposts) that are missing are harder to spot. Incompleteness of attribute data is easy to spot because of missing values result in empty fields. In theory an empty data field can be correct so which values can occur and could or should be empty (Null) should be defined by forehand. In general, preventing null-values should be pursued.

Temporal accuracy

Part of the definition of NI is that it is generated real time, so updating information plays a critical role in the process of creating NI. A spatial dataset is a conceptualization of the real world, which is constantly changing. Badard and Richard (2001 cited in Halls, 2001, p.2) "note that ... the issue of how to convey and incorporate update information into the spatial database has become critical". Both the positional and attribute data need to be up to date. For instance outdated information about storage locations of hazardous substances (e.g. toxic, flammable or explosive substances) can pose a threat to public safety in case of a disaster.

"Every spatial object used in a GIS has a temporal validity as well as one or many attribute values" (Ott and Swiaczny, 2001, p.55). Especially when using multiple sources to create NI, overall data consistency is important. Information is only valid for a certain amount of time (figure 4).

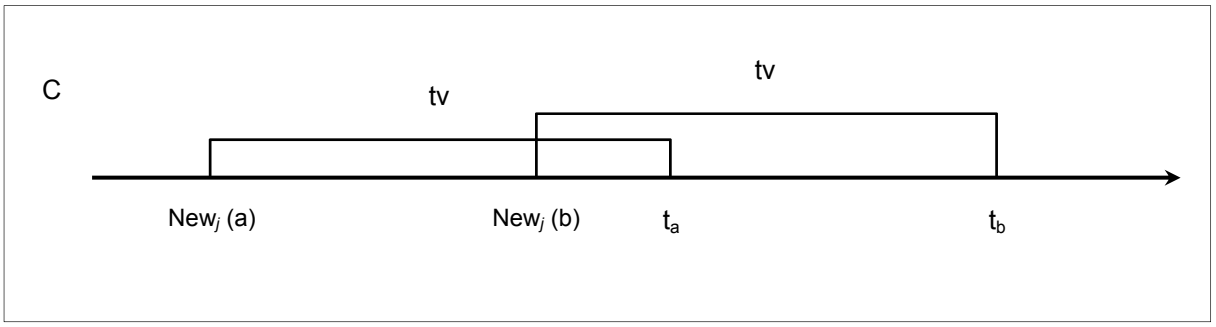


Figure 4: New$_j$ (a) is the moment on which component C within a data set is updated. The temporal validity tv of component C expires at time t_a. So the next update new$_j$ (b) should to occur before time t_a for component C to retain temporal validity (from: Anderson and Filipe, 2003, p.180).

The temporal validity of (objects within) a data set can vary when it is differently applicated. For this, users have to be able to determine the temporal validity, which is dependent on the time of the last update and the rate of change. The rate of change provides an estimate of the rate at which objects that are displayed in a data set change. The higher the rate of change, the faster the timeliness of the data set deteriorates. Another interesting element is how long it takes for a change in the real world is reflected in the data set (temporal lapse).

It may be assumed that most users of a NI product do not have the means to check the different elements of quality, due to lack of necessary hardware, software, staff or knowledge. They have to rely on the checks the producer of the NI product performs.

2.2.3 Reliability

Many users have developed a certain level of dependency towards NI products, and not being able to 'consume' the NI product when needed is ranging from annoying to catastrophic. For public users the individual impact in most cases is small, e.g. being late. For corporate and governmental users the stakes are higher. Depending on the business, it means loss of reputation, loss of money or even loss of life. So reliability of NI means users can depend on it and that it is accessible when needed.
Two main risks are repeatedly mentioned in literature and on the internet, i.e. server downtime and process discontinuity. These apply to both the producer of NI and the supplier of raw data, because if the raw data is not delivered to the producer of NI, the product cannot be created.

Server downtime can be planned or unplanned and have different causes, e.g. server maintenance, failures in hardware or software, fire and natural disasters (table 3).

Causes of downtime	Examples
Planned administrative downtime	Upgrades for hardware components, firmware, drivers, operating system, and software applications.
Component failures	Faulty server components, such as memory chips, fans, system boards, and power supplies. Faulty storage subsystem components, such as failed disk drives and disk controllers. Faulty network components, such as routers and network cabling.
Software defects or failures	Drive stops responding, operating system stops responding or reboots, viruses, or file corruption.
Operator error or malicious users	Accidental or intentional file deletion, unskilled operation, or experimentation.
System outages or maintenance	Software or systems requiring reboot, or system board failure.
Local disaster	Fires, storms, and other localized disasters.
Regional disaster	Earthquakes, hurricanes, floods, and other regional disasters.

Table 3: Causes and examples of downtime (from: Microsoft, 2013).

Related to availability, Kaufmann and Hintergräber (2012, p.18) state that the service should be available for a certain amount of time a year (in percentage) in combination "with a maximum allowable downtime (MADT) of the service, expressed in hours per disturbance". Related are the recovery point objective (RPO) and the recovery time objective (RTO), which according to the Business Continuity Management Institute (n.d.) are two fundamental terms. RPO is defined as "the maximum acceptable level of data loss following an unplanned event" and RTO as "the length of time that a business process could be unavailable before the business unit's operations are significantly impaired" (Business Continuity Management Institute, n.d.). The RTO is the period between the disruption and the resumption of the service. The maximum tolerable period of down time (MTPOD) starts at the same time as RTO, but continues until all business processes should be fully restored. If not, an organization's viability is at serious risk (figure 5).

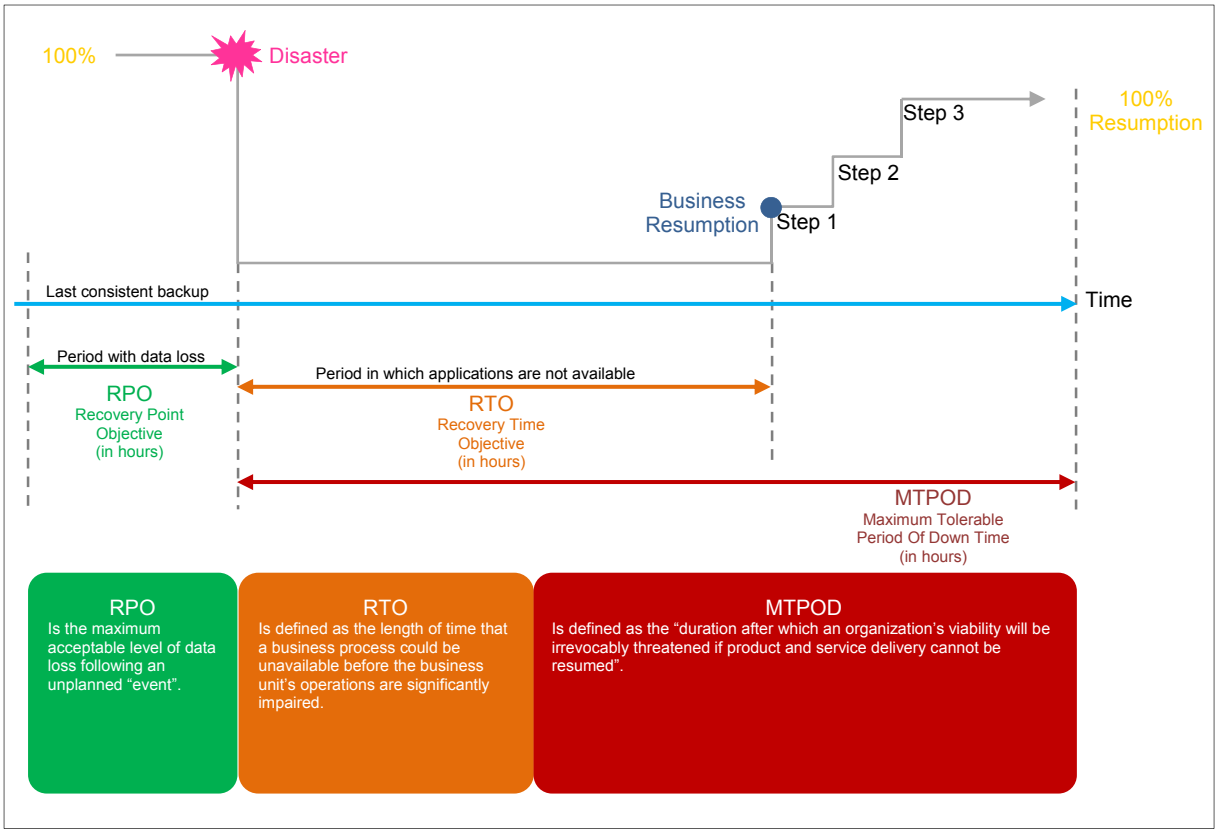


Figure 5: Visualization of recovery point objective (RPO), recovery time objective (RTO) and maximum tolerable period of down time (MTPOD) (from: Business Continuity Management Institute, n.d.).

Off course, it is better to prevent than to cure. But in case of server downtime the organization has to "document procedures (including necessary arrangements) to ensure continuity of activities and management of a disruptive incident" (St-Germain, Aliu, Lachapelle and Dewez, 2012, p.6). Information about actual occurred server downtime may well be regarded as sensitive and companies might not be inclined to share this kind of information. Second best are the theoretical guaranteed values specified in the SLA.

Process continuity concerns a reliable data flow from supplier to producer, data processing and delivery of NI to the client. The foundation of a reliable, continues data delivery process are agreements and guarantees with suppliers and should be recorded in a contract. This is something that a company can exert influence on.

For data processing, a company should not be dependent on one employee. Critical knowledge about (part of) the process should be documented and shared. To ensure NI delivery, arrangements for raw data, production process and the NI should be taken. A solution could be software and data escrow. Escrow is a service offered by a natural third party, where that party keeps a back-up, in this case of data and software. In case of unexpected events (e.g. server downtime or fire), the third party ensures provision of the service (figure 6). Business bankruptcy has an even bigger impact in that the delivery would stop completely. Because of financial reasons and the length of the period, escrow usually does not suffice.

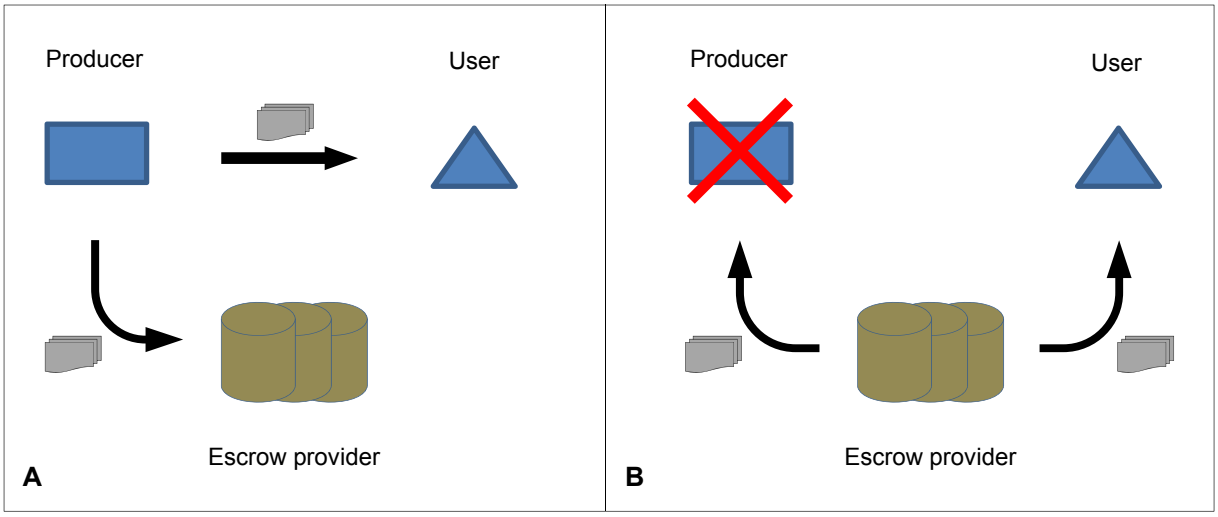


Figure 6: Schematic display of escrow service. A: The producer delivers information to both user and escrow provider. This ensures information back-up. B: The escrow service takes over delivery of information when for any reason the producer cannot. Information is also send back to the producer so the producer is able to restore its own system. This arrangement does not work when the producer goes bankrupt.

Information with a geo component is often bigger than information without it, and thereby the challenge to get the information to the client as well. Especially when the client uses a mobile device. The OGC is developing a new standard, GeoPackage, that is supposed to be the new universal geodata file format. The standard “provides an open, non-proprietary, platform-independent container for distribution and direct use of all kinds of geospatial data” (Open Geospatial Consortium, 2013) and should, among other advantages, be a big help for mobile device users who require geospatial application services … and who operate in disconnected or limited network connectivity environments” (Open Geospatial Consortium, 2013).

2.2.4 Security

“Information resources are critical assets of most 21st century businesses and the protection of these assets is as crucial as the protection of physical assets” (Dzazali, Sulaiman and Zolait, 2009, p.584). Information security applies to all kinds of information, ranging from private to business secrets to national security.

Threads on information security are not only external (e.g. corporate or government espionage), but also internal (figure 7). Logan and Clarkson (2005) even wrote a paper on “overcoming the insider threat”. They found that spending money alone is not enough, but that adopting “a simple method of insider risk assessment that targets the interactions of people, processes and technology ... will improve the likelihood that insider vulnerabilities will be exposed” (Logan and Clarkson, 2005, p.215).

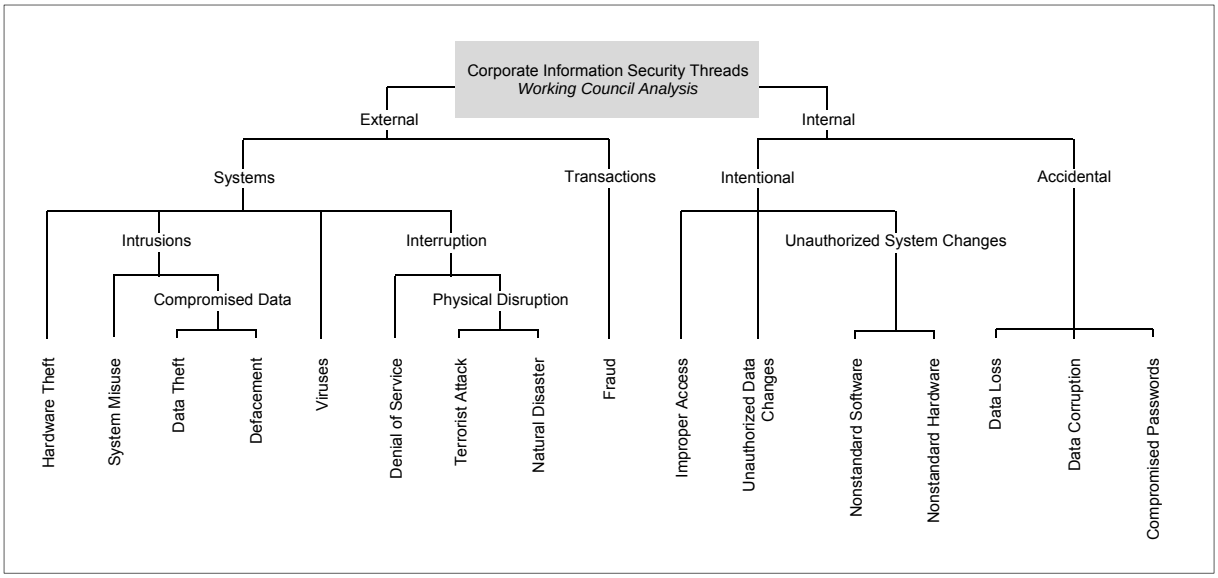
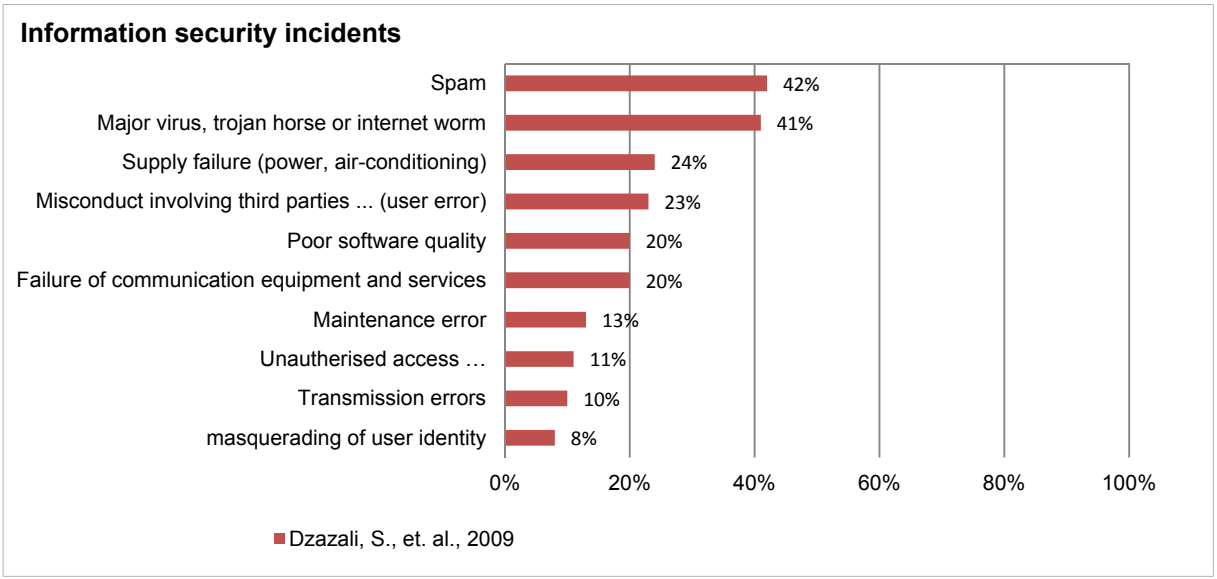


Figure 7: Spectrum of external and internal security threats (from: Working Council for Chief Information Officers, 2003, p.5).

An important measure is to inform the individuals in an organization about their role in information security through awareness campaigns. Examples are the awareness posters of PeopleSoft's "Operation Barbed Wire" (figure 8) and the mandatory training on the secure usage of laptops for all Bank of Scotland staff with laptops (Working Council for Chief Information Officers, 2003, p.31).

Figure 8: Examples of awareness posters from PeopleSoft's "Operation Barbed Wire" (from: Working Council for Chief Information Officers, 2003, p.33).

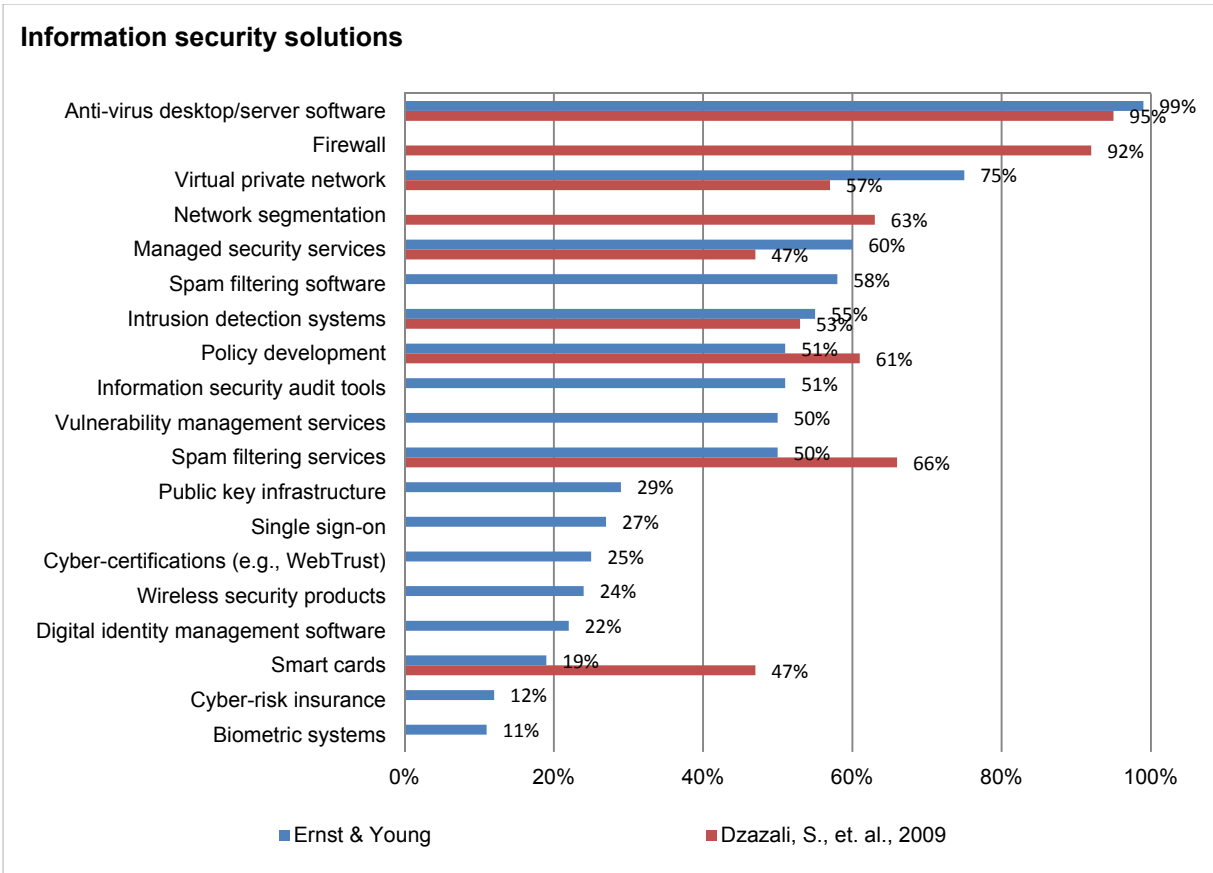
Dzazali, Sulaiman and Zolait (2009, p.590) provide an overview of the top ten incidents experienced by organizations that participated in their research, with spamming and 'malicious code' (e.g. virus or worm) occurring most frequent (graph 2).



Graph 2: Information security incidents (based on information from: Dzazali, Sulaiman and Zolait, 2009, p.589).

The number of threads are increasing, and still more than two thirds (69%) of the participants of the Ernst & Young 2012 Global Information Security Survey have not implemented a formal security architecture framework (e.g. The Open Group Architecture Framework or ANSI/IEEE 1471:ISO/IEC 42010) (Ernst and Young, 2012, p.23). This lack of preventative measures is often due to financial reasons because it is not seen as part of the business goals (Saleh, 2011. p.316).

Dzazali, Sulaiman and Zolait (2009, p.591) found that the top three measures that organizations deploy are installing anti-virus and firewall software and access control to the information system. The global research of Ernst and Young (2004, p.22) shows similar results, with slightly more emphasis on more comprehensive measures (graph 3). The similarity in the findings of both researches, would imply that in five years not much has changed in information security. According to Ernst and Young organizations are struggling to keep up with new ICT and social developments. Despite the effort companies and governments put in improving information security, "they continue to fall behind, creating an information security gap that grows ever larger" (Ernst and Young, 2012, p.5).



Graph 3: Information security solutions (based on information from: Dzazali, Sulaiman and Zolait, 2009, p.590 and Ernst and Young, 2004, p22).

An organization should be doing a structural effort, supported by management, to prevent security breaches by active development of security management (figure 9).

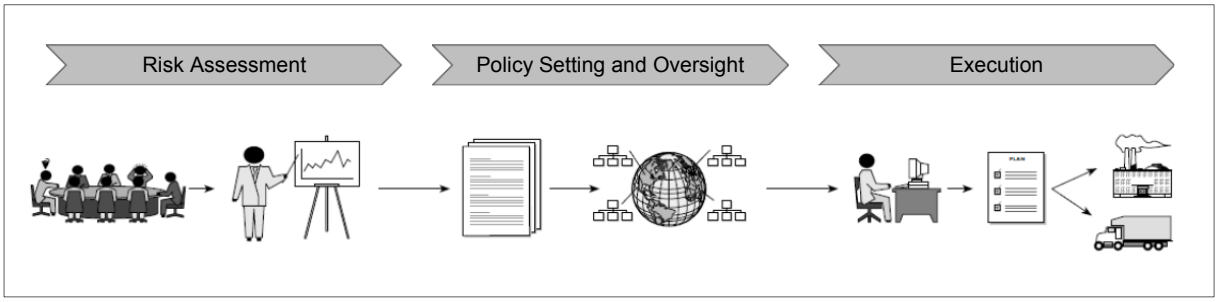


Figure 9: Business protection life cycle (from: Working Council for Chief Information Officers, 2003, p.9).

2.2.5 Transparency

For the client to be able to evaluate the usability of NI, he needs to have insight in the producers business information (e.g. production process, security management or reliability of servers). Thus evaluation of the usability of NI, stands or falls with the cooperation of the producer.

Some information might not be released at all and when it is given, the biggest question is whether the information is reliable. Companies might not be inclined to share information that could harm their business or reputation and it is difficult if not impossible to find out. So for some points (e.g. server

downtime), the investigator has to be satisfied with theoretical guaranteed data instead of real life figures.
Because of the dependability of the user upon the producer to release the needed information, be it via a survey or website, transparency deserves to be identified as a separate aspect. It will be rated by the amount of information provided.

2.3 Setting the standard

The aspects that influence usability of NI and in what way they do is covered in the previous paragraph and the result is presented in table 4. In the next chapters, this theory is applied in practice by comparing five NI products with each other in a benchmark wherein the standard is based upon the previous findings.

Aspects	Indicators	methods
Privacy	confidentiality	Does the personal data protection act apply?
		Is it being followed?
Quality	Correctness	Is the lineage completely documented?
		If no, what parts of lineage are documented?
		Are the attribute data checked?
		If yes, how is it checked (e.g. cross referenced with other data sources)
		Is the relative positional accuracy checked?
		If yes, how is it checked?
		Is the absolute positional accuracy checked?
		If yes, how is it checked?
	Completeness	Attribute data: is the existence of invalid empty fields checked?
		Objects: is object completeness checked against agreed data model?
	Temporal accuracy	What is the update frequency of the data in the data set?
		What is the rate of change of the data in the data set?
		How long does it take for a change in the real world to be reflected in the data set (temporal lapse)?
Reliability	Availability (down time server)	What is the guaranteed server uptime?
		What is the achieved server uptime (in the past year)?
	Continuity (in case bankruptcy delivery or producer)	Have measures been taken in the case of bankruptcy?
	Arrangements	Are terms of use documented?
	Black box (producer)	Is the process secured for the producer?
Security	Active security management	Is an active security management applied?
		Is it supported by management?
		Are information security risks measured and prioritized?
		Have security measures been taken?
		Do security measures apply to a standard (e.g. The Open Group Architecture Framework and ANSI/IEEE 1471:ISO/IEC 42010)?
	Protection against malicious software (e.g. virus)	What measures have been taken?
	Protect against use by unauthorized person	How is users' access to information controlled?
Transparency	Openness	Is the appropriate information released?

Table 4: List of aspects related to the usability of NI (layout is based on table 1 from: Daas, Ossen, Vis-Visschers and Arends-Tóth, 2009 and Daas, Ossen and Arends-Tóth, 2009).

CHAPTER 3
Research methodology

3.1 Introduction

This thesis is a study to the usability of Networked Information and how this might be measured. The underlying idea is that if the usability of the NI product can be measured, and the aspects that influence it are known, than producers of NI will be motivated to improve the usability of their product. For this, the following main question is formulated:

Which aspects of networked information influence its usability and can the aspects be used to measure and compare the usability of networked information products?

The main question is divided in four sub-questions:

1. What is usability in relation to NI?
2. What requirements does NI need to have to be fit for use?
3. How can the requirements for the fitness-for-purpose of NI be measured?
4. To what extent can the producer of NI positively influence these aspects?

To be able to answer the second part of the main research question (can the aspects be used to measure and compare the usability of networked information products) and the third sub-question, a questionnaire was used for collecting data and a benchmark was used to compare the different products using the collected data.

In practise, this exercise will be used by potential clients who need a NI product for a specific purpose. They will want to objectively and equally compare multiple products that can serve the intended purpose and choose the one that best suits their needs.

A benchmark is very suited and much used device for comparing products. It provides a structure in which the different products are compared to ‘the best in the market’. Instead of the best in the market, the users’ needs or demands are translated to a baseline that is used to compare the products to.

Maybe the biggest challenge is to get comparable and above all trustworthy data. This is why the questionnaire has a mix of closed and open questions. The closed questions provide information about the different products that is structured, easy to compare and thus suited for use in a benchmark. This facilitates an objective comparison of the different NI products. The open questions are used as follow-up questions to (a clusters of) the closed questions. The purpose of the open questions is twofold. First to be able to check the answers, at least to some degree and second, to give the participant the opportunity to clarify the answer.

3.2 Review of literature study

The literature study is used to define NI, usability of NI and to find the aspects of NI with which the usability of a NI product can be evaluated (chapter 2.2). This is a theoretical approach to 'usability' of NI instead of doing (qualitative) research to what people understand it to be.
The findings have been used to create a questionnaire to measure the usability of a number of NI products and to compare the results by means of a benchmark.

Keywords: Networked Information, Service oriented architectures (SOA), Information as a Service (IaaS), Data as a Service (DaaS), System as a Service (SaaS), Model as a Service (MaaS), Usability, Privacy, Quality, Reliability, Security, Transparency

3.3 Data collection

3.3.1 Introduction

Because this is a research about information as a product and the degree to which it can be used for a specific purpose, based on facts and measurement instead of emotion or a deeper meaning, the research method is quantitative. For instance, there is no opinion or faith about whether or not NI is available or legal, it either is or it is not (although in respect to the latter, high priced lawyers could disagree). Data about the NI products is accumulated through a combination of web research and a questionnaire. The products are compared using a benchmark.

3.3.2 Questionnaire

The information that is necessary for this research is not available from public sources. Therefore the questionnaire is the main source of information for the benchmark. In comparison to solely performing web research, it has the advantage that all participants have an equal opportunity to submit information.

A lot of information is available on how to set up a questionnaire (formulation and placement of questions, layout and testing the questionnaire), how to engage participants and how to manage the data (e.g. Bradburn, Sudman and Wansink, 2004; Scheuren, 2004; Trochim, 2006 and Gray, 2004).

The questionnaire is divided in three parts and consists of 27 questions and a number of follow-up questions. In the first part the participant is instructed and terms used in the questionnaire are defined. Part two contains easy and neutral product descriptive questions. The answers given in this part help to describe the product in chapter 4 of this thesis. The questions in part three deal with the aspects that influence the usability of the NI (chapter 2). This part consists of closed questions, with categorical response choices, and open follow-up questions.

The highly structured level of the closed questions make the results very suitable for quantitative analysis (Gray, 2004, p.215) in a benchmark. The function of the open questions is to verify the given answers as far as possible and to allow the participant to clarify his answers. The answers to the open questions are not directly used in the benchmark.

Efforts have been made to construct the questions as neutral and as clear as possible. To ensure the quality of the questionnaire, all questions (table 5) and answer possibilities to the closed questions (table 6) have been measured against a number of guidelines.

The questions are:
valid (they measure what they are intended to measure)
reliable (they provide consistent measures in comparable situations)
clear and use simple language
concise and specific
possible to answer and relevant to the respondent
avoiding biased terms
not using negatives
consisting of only 1 part (not “double-barrel”)
categorised

Table 5: Guidelines for question format (adapted from: Welchel, 2003; Gray, 2004 and Scheuren, 2004).

The answers are:
mutually exclusive (can select only one appropriate answer)
exhaustive (all possible answers are listed)

Table 6: Guidelines for answer format (adapted from: Welchel, 2003; Gray, 2004 and Scheuren, 2004).

The questionnaire was submitted to a small sample of the population and conducted face-to-face by an interviewer.

The subject (the population) of this study is networked information (NI), which is generated real-time from at least one source available via the internet and it is distributed online. It is information which is the result of a Model as a service (MaaS). In a short amount of time, the number of products has increased dramatically. In the Netherlands alone, the number of NI is probably in the thousands and the use varies from private and free of charge (e.g. weather prediction or information about public transport) to corporate and paid (e.g. assisting in disaster management). Governmental organizations are big (probably the biggest) producers of data and present a large part of that data as a service (e.g. through PDOK). They also create a lot of NI (e.g. maps that inform citizens about public facilities and geo statistics). Companies create their own products, aimed at supporting business operations of

other organizations or at individuals. Especially this last category is enormous and rapidly growing. Products are distributed via website or apps for mobile devices and the relative ease with which the products can be developed in combination with the amount of the potential users, this is an attractive marked.

The sample used in this study is randomly selected from the population, with the exception that no two selected products serve the same purpose. The hope is that participants (being the producers of the NI) are more inclined to provide information when their direct competitors do not participate in this research.
The sample consists of five NI products and two data as a service (DaaS), which can be used as an input for NI. The DaaS are not part of the population, but share many of the NI properties. They are included to see if the results show significant difference in comparison to NI.
Compared to the size of the population, the sample is very small. Because the main goal of this thesis is to identify aspects that influence the usability of NI and show if it is possible to use these aspects to measure the usability of NI, instead of for instance researching the state of the usability of NI in the Netherlands in 2013, which would mean testing of a much larger sample, the sample size used in this research is justifiable.

The questionnaire is pre-tested by two parties, being RTL Nederland and BridGis. Any remarks have been discussed and incorporated. The final version of the questionnaire is included in appendix A.

The first contact with the participants was by telephone, explaining the research and inviting them to participate in the survey. In case of positive reaction, a date for the interview was set and a conformation mail, including the interview, was send to the participant. The interview itself was held as soon as possible, so the participant had enough time to prepare and consult documents when needed and the topic was still in focus.

The main reason to conduct the questionnaire face-to-face was to increase the response rate. One of the biggest drawbacks of an interviewer present is the potential influence (consciously or unconsciously) on the participant and with that, on the outcome of the questionnaire. To keep this to a minimum, there was as little as possible interaction between interviewer and participant, all participants got the same questions and were treated in the same way. The interviews were all held at the participants location.
A consequence of interviewing in person is that the participants were looked for in the Netherlands because of travel time and expense.

3.3.3 Publicly available product information

What information about a NI product is publicly available is determined through web research. The purpose of this method is twofold. First it is used to verify answers given in the survey (as far as possible). Second, and more important, the amount of information available determines part of the

final score of the aspect ‘transparency’. The web research was conducted before the interview was held.

A vendors website has to give some basic information about the product on the basis of which a customer can make an initial estimate whether the product is suitable for the intended use. In view of the aspect ‘transparency’ a minimum amount of information is expected to be found (table 7). The extent to which the information is actually offered on the website is of importance, not the ease with which the information is found.

Information	Motivation
Contact information for questions about the product	Potential clients have to be able to ask their questions.
Used sources (at least the main sources) in creating the product	For getting insight in the structure of the product.
Intended use of the product	For assessing whether the product meets the needs of the customer.
Costs of using the product	For assessing if use of the product fits in the available budget.
Manual (short) on how to use the product	For assessing if the client is able to use the product (e.g. appropriate hard- and software and knowledge).
References from (satisfied) customers	For making inquiries about customer satisfaction.

Table 7: Minimal expected amount of information on the vendors website about the product and the motivation of why it is expected.

3.4 Data Analysis

The gathered data was used to compare seven products from different vendors (five NI and two DaaS) using a benchmark. The products have been compared to each other and to a theoretical standard (baseline) instead of the product recognised as the best within the area.
The baseline that is used here (Appendix B) is a working example. For every user and every use, the focus and therewith the values for the different aspects, can be different. For some aspects, like privacy, for which the producer has to comply with the law, the baseline will always be the same. On the other hand, privacy, when looking at the way data is used in the production process of the NI, is not always an issue. For instance no personal related information is used with flood prediction.

The results of the questionnaire are presented in a table. Table 8 shows an example of how the data, gathered through questionnaire and web research, are presented. The scores are expressed in terms of + (above baseline), o (on baseline) and - (below baseline). The values for each aspect are transformed to a score by dividing the number of questions the product matched (o) or exceeded (+)

the baseline by the number of questions. There has been made no difference in scoring if the product matched (o) or exceeded (+) the baseline. Both is positive. For example, product A matched the baseline three times out of four questions for Quality. That gives a score of 3/4 (3 positive scores divided by 4 questions in total).

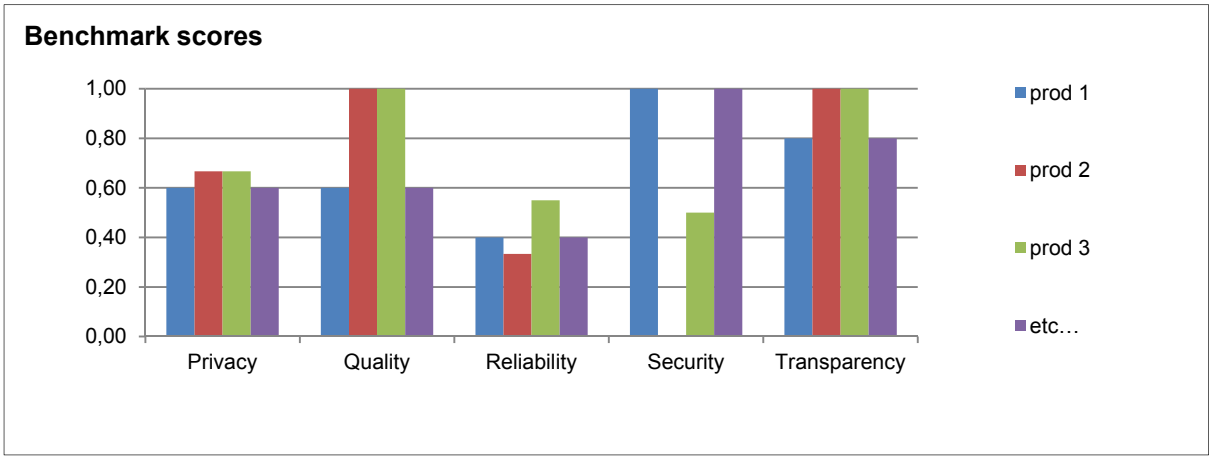
The total score, which indicates the usability of each product, according to the baseline that is used in this study, is generated in the same way, namely by dividing the number of questions the product matched (o) or exceeded (+) the baseline by the number of questions. In generating the total score, all questions weigh the same.

Different NI products →	Prod A	Prod B	Prod C	Prod D	Prod E
Quality	**3/4**	**2/4**	**3/4**	**2/4**	**3/4**
Q10	o	-	+	-	+
Q11	o	o	-	-	o
Q12	o	-	o	o	o
Q13	-	o	o	o	-
…	…	…	…	…	…

↑ Aspects ↑ Scores

Table 8: Example of a benchmark score list. The scores are derived from the number of times the product matched (o) or exceeded (+) the baseline, compared to the number of questions. There is no difference in scoring if the product matched (o) or exceeded (+) the baseline, both is positive.

A bar chard is then used to visualize the scored percentages of each product per aspect (graph 4). The percentages are generated using the obtained score. For example, a score of 3/4 is 75% and a score of 1/2 is 50%.



Graph 4: Example of a bar chard that presents the scores of each product summed per aspect.

The bar chard provides a good overview of the obtained scores for each aspect. This could show a trend, for instance, all products score bad at reliability, or only the producers of free products score very good at transparency.

Because there is no other way to verify the information that was supplied by the producers following the questionnaire, customers of those producers have been approached when possible. They have been asked about their experiences with the product, regarding the aspects that influence the usability. This step lies beyond the original scope of the research, but it is interesting to see if the image the producer outlines corresponds to that of the customer. The customers have been contacted by phone.

CHAPTER 4
Results

4.1 Introduction

For this part of the research, seven products have been examined for which eight parties were interviewed. The products have been compared to each other and a baseline. The questionnaire that is used to collect the data, is divided in three parts and consists of 27 questions and a number of follow-up questions. The questions in part three deal with the aspects that influence the usability of the NI (chapter 2). This part consists of closed questions, with categorical response choices, and open follow-up questions.
The highly structured level of the closed questions make the results very suitable for quantitative analysis (Gray, 2004, p.215) in a benchmark. The function of the open questions is to verify the given answers as far as possible and to allow the participant to clarify his answers.

The participating organizations are located throughout the Netherlands and of those organizations, four are commercial companies and three are governmental agencies (figure 10).

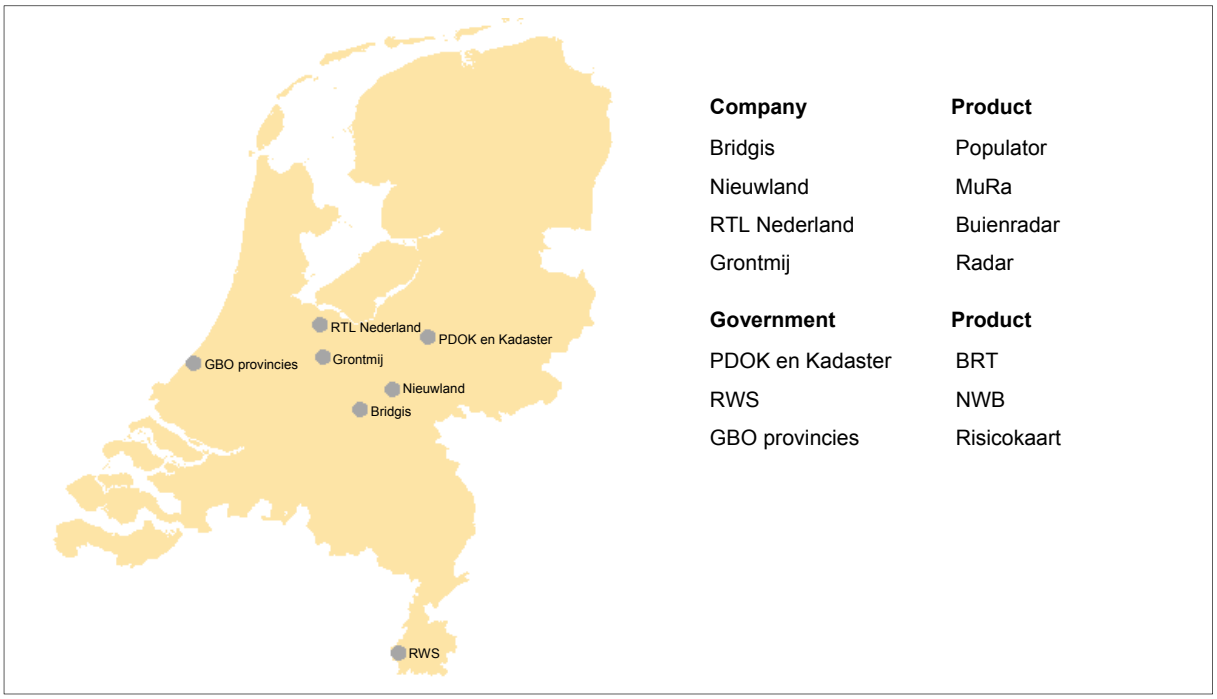


Figure 10: The approximate location of the participating organizations in the Netherlands.

Five of the products are NI as defined in chapter 1 of this thesis and two products (BRT and NWB) are Data as a Service (DaaS). A number of other characteristics are shown in table 9.

	Populator	MuRa	Buienradar	Radar	BRT	NWB	Risicokaart
Type	NI	NI	NI	NI	DaaS	DaaS	NI
Use	paid	paid	free	paid	free	free	free
Users (No.)	100s	10s	10.000s	10s	10.000s	1000s	10.000s
Organization	company	company	company	company	government	government	government
Influence	safety	financial	social	financial	financial	financial	safety

Table 9: Characteristics of the participating products. The characteristic 'influence' is based on the use of the product. The influence of Populator is safety related, because it is used for i.a. creating evacuation plans. Risicokaart is also safety related, because it helps e.g. emergency services do their jobs. MuRa, Radar, BRT and NWB are used to in the business of organizations and help improve operations and thus have financial impact. Buienradar is rated social, because it is mainly used to help individuals organize their daily outdoor activities.

The interviews were conducted within a few weeks' time. All participants were very cooperating and have taken plenty of time to speak to me. Also most participants were well prepared. The interview started with a short introduction about the research, the time necessary for the interview and what would happen with the results.

After all interviews were held, the individual results, consisting of the written summary of the interview and the individual scores, as included in sections 4.3.1 to 4.3.7, were send to each of the participants, to give the opportunity to correct any inconsistencies before the results are published.

All participants replied with minor remarks, ranging from improvements on style to adjustments in the text. In response to a question about the way in which the scores were calculated, RWS was contacted by phone.

For only three of the seven products, users have been contacted personally to comment on their experiences. In case of one product, the internet was used to find user experiences, in another case, an internal customer satisfaction survey was used. Two organizations had rather not I contacted their clients because of business developments. This is respected and for these products, users have not been contacted.

4.2 Summarized results

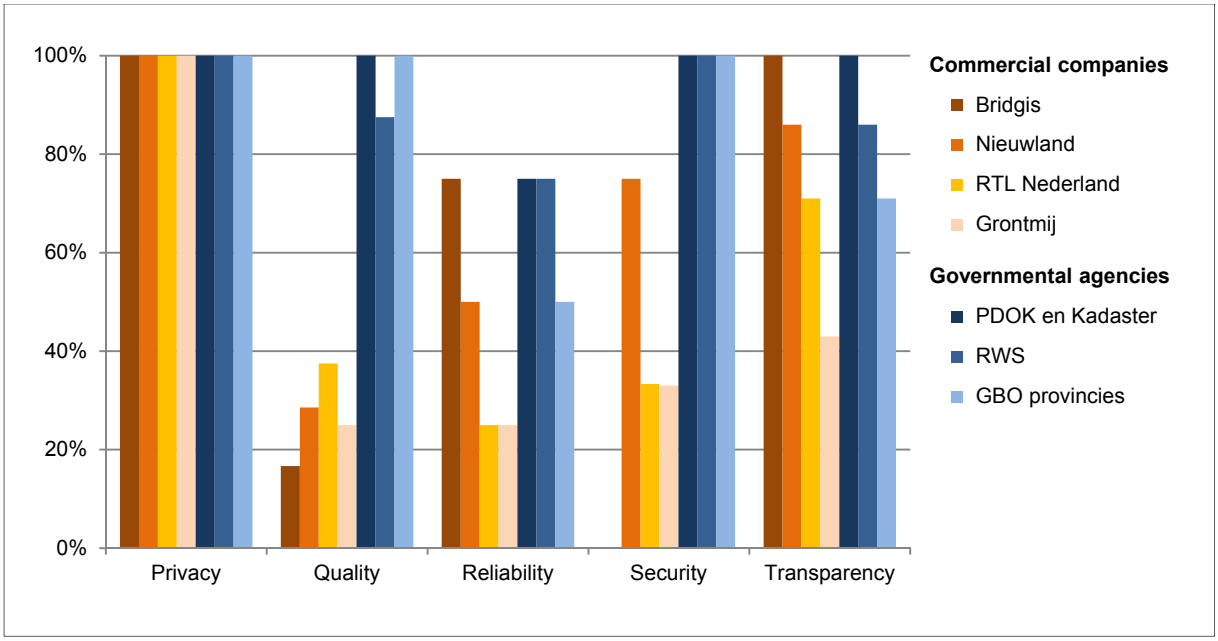
Based on the results of the interviews, the products are compared to a working standard (appendix B) and each other. The results of all interviews as well as the scores for transparency, are listed in table 10 and graph 5. Looking at the table and the graph, a number of things stand out.

The product characteristic "organization" (table 9) seems to be of influence on the usability of NI products. According to the total scores, the products of the government agencies are much more usable than the NI of the companies. Especially in the field of data quality and security, government agencies stand out. No apparent relations between the other product characteristics from table 9 and the aspects of usability are found.

Categorised questions	Score						
	Bridgis	**Nieuwland**	**RTL Nederland**	**Grontmij**	**PDOK en Kadaster**	**RWS**	**GBO provincies**
Privacy	**3/3***	**3/3***	**3/3***	**3/3***	**3/3***	**3/3***	**3/3***
9	n/a	n/a	n/a	n/a	n/a	n/a	n/a
10	o	n/a	n/a	n/a	n/a	n/a	n/a
11	o	o	o	o	o	o	o
Quality	**1/8**	**2/8**	**3/8**	**2/8**	**8/8**	**7/8**	**8/8**
12	-	o	+	o	+	+	o
13	n/a	-	-	-	o	o	o
14	-	-	o	-	+	+	+
15	n/a	n/a	-	-	+	-	+
16	o	-	o	o	o	o	o
17	-	-	-	-	o	o	o
18	-	-	-	-	o	o	o
19	-	o	-	-	+	+	+
Reliability	**3/4**	**2/4**	**1/4**	**1/4**	**3/4**	**3/4**	**2/4**
20	-	+	+	o	o	+	+
21	o	-	-	-	-	-	-
22	+	o	-	-	o	o	-
23	o	-	-	-	o	o	o
Security	**0/4**	**3/4**	**1/4**	**1/4**	**4/4**	**4/4**	**4/4**
24	-	o	-	-	+	+	+
25	-	-	-	-	o	o	o
26	n/a	+	n/a	n/a	+	+	+
27	-	o	o	o	o	o	o
Transparency	**7/7**	**6/7**	**5/7**	**3/7**	**7/7**	**6/7**	**5/7**
Interview	o	o	o	o	o	o	o
Contact	o	+	o	-	o	o	+
Used sources	+	+	o	-	o	o	+
Intended use	o	o	o	o	o	o	o
User costs	o	-	o	-	o	o	o
Manual	+	o	-	o	+	+	-
References	o	o	-	-	o	-	-
Total score	**14/26**	**16/26**	**13/26**	**10/26**	**25/26**	**23/26**	**22/26**

Table 10: Results of all interviews combined with the scores for transparency. The numbers in the first column correspond with the interview questions (appendix A). The answers have been rated according to the baseline (appendix B). The score that is given for each aspect, compares the number of questions that were answered on (o) or above (+) the baseline, to the total number of questions.

*Score 'o' for question eleven indicates the NI product fully complies to the Dutch personal data protection act and thus scores 3/3, no matter the scores for questions nine and ten.



Graph 5: Scores per category and per company. Organizations with shades of red are commercial companies, organizations with shades of blue are governmental agencies.

All participants score equal at privacy. Bridgis is the only organization that uses a data product in which personal information is incorporated, to create their product. All other organizations do not use personal information and thus fully comply to the Dutch personal data protection act. Therefore, unlike the standard scoring method, all organizations score three out of three.

What particularly stands out about quality, is that without exception, governmental agencies score significantly higher than the companies. Almost all organizations check the quality of the data, but to what extent and frequency varies widely. Companies mostly check one of the three aspects of quality (correctness, completeness and temporal accuracy), which is partly explained by the fact that some data that is part of NI is being imported by the client who then is automatically responsible for the quality of that data.

Scores on reliability are very divers and do not have a noticeable connection with any aspects like size and type of organization, whether the product is free or paid or the number or type of users of the product. The only two items most organizations have in common is that knowledge about the production process is well known within the organization (question 20) and that hardly any organization have taken measures to ensure the continuation of the service, would the organization go bankrupt (question 21). The latter does not really apply to government agencies, because they are very unlikely to abruptly cease to exist.

Regarding security, organizations that pay most attention to the security of information are the companies that have had unpleasant experiences in this area and all governmental agencies. But, according to the interviews, the focus mostly lies on external security and less on internal security.

On transparency, results of the interview show that most organizations, be it governmental or commercial, share the view that it is important to inform (potential) users. It is the best way to get attention to a product and to get new customers.

On the basis that in this study all questions weigh the same, and compared to the working example baseline, the service (DaaS) of PDOK in combination with the data of Kadaster has the highest score (twenty-five out of twenty-six) of all products. Looking solely at the NI products, Risicokaart of GBO provincies (twenty-two out of twenty-six) would be the most usable product.

4.3 Results by product

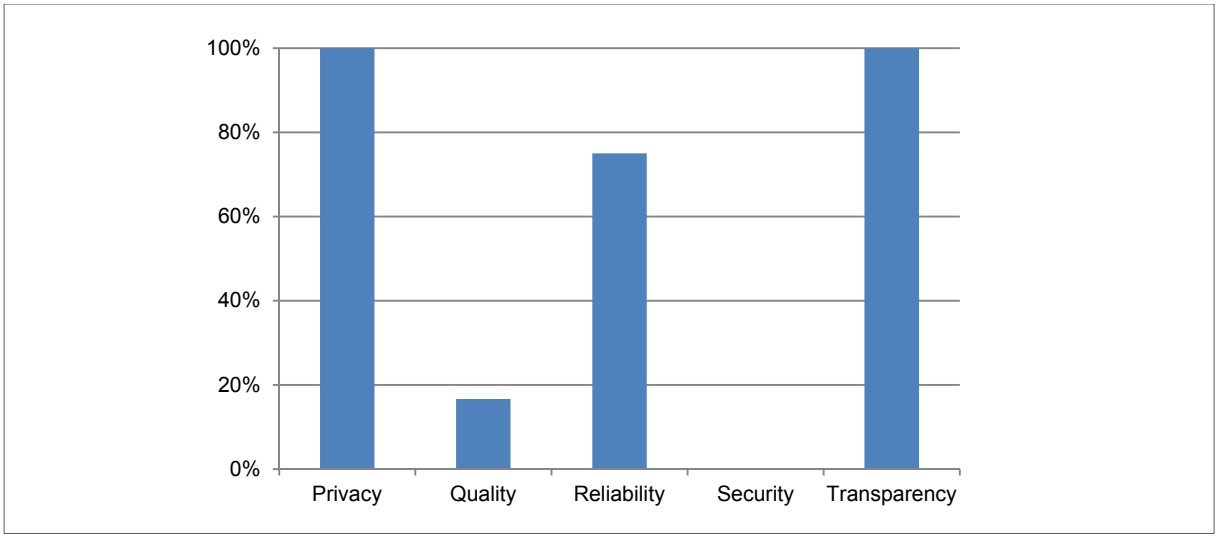
In this chapter, every company and product is introduced and the interview results for each product are summarized and commented. Only two products are DaaS (NWB and BRT), all other products are NI as defined in this research.

4.3.1 Populator by Bridgis

Bridgis is a small to medium sized company that is active for over 20 years on the market for geographical information. It aims on the development, combination and distribution of geographical information. One of the products Bridgis has developed and sells is the Populator.

The Populator is constructed to provide uniform population data. It is an online information service that calculates the potential population present for any specified area at any given time. The data can be used in the calculation of risks related to e.g. facilities, pipelines and transport of dangerous goods. It also supports the preparation of a regional risk profile and provides necessary information in case of an disasters like flooding or fire (Bridgis, 2013).
An example of the use of the Populator is project Reality Check, in which an evacuation during a flood is simulated. The Populator data is used to calculate the amount of people and the type of people that are located in a specific area. The project is initialised by the consortium Flood Control 2015, which aims for “a really substantial improvement in operational flood protection worldwide” (Flood control 2015, 2012).
In the future, the use of the information that is provided by the Populator will be more diverse (for instance for marketing). One of the developments could be adding more diverse datasets.
To get information, the user (graphically) defines the area of interest, the population types and selects the maximum population or determine the time frames.
Bridgis has a contract with the Dutch government (ministry of Infrastructure and Environment) in which the use by lower governments (municipalities, provinces and safety regions) is regulated. User costs are calculated on the basis of a fixed subscription fee, to which the ministry of Infrastructure and Environment contributes a fixed part.

The interview with Bridgis about the Populator was held on the twelfth of June 2013, from 13.00 to 14.00 PM, at their office in Tiel. Two persons, a sales manager and a database manager, were present. The individual results of Bridgis are combined in graph 6.



Graph 6: Combined results after the interview with Bridgis and the web research.

Bridgis is the only organization in this research that makes (indirect) use of personal information in its product, through at least one of their data suppliers. The data is gathered by the supplier by interviewing individuals. The outcome is than combined with other interview results and aggregated to the level of six digit postal code. After this operation, the data can no longer be traced to a person and is therefore compliant with the Dutch personal data protection act.

Bridgis scores lowest on quality. The data model is complete, but they have not implemented quality management, do not perform data quality checks and have only partially documented data lineage and metadata. The participants of the interview indicated that the company has started developing quality management, including data checks (e.g. sanity and cross checks).

Bridgis scores best on reliability. It is the only organization that offers clients the security that the service will stay available for half a year in case Bridgis would go bankrupt. Also Bridgis guarantees and delivers their clients the highest server uptime. In the last three years they have had no unscheduled server downtime. The only drawback is that only one person knows and understands the full production process of the Populator. Although the process is completely documented, in case of prolonged absence, for whatever reason, of that employee, the continuity of the service could be threatened.

Bridgis scored no points for security, which doesn't mean they ignore the issue completely. Because the company and the product are relative small, Bridgis does not expect to need an extensive security architecture. Possible security risks have not been assessed and documented, but with common sense and regular updates of e.g. anti-virus software and firewall, Bridgis expects to have covered any real threats. Also, Bridgis is a small company, employees know each other and have a strong connection and loyalty to the company. This makes the risk of internal security threads small.

Bridgis has the maximum score on transparency. Before, during and after the interview, Bridgis was very cooperating and willing to provide information. The information about the product that can be gathered through the website of Bridgis is very satisfying as well. Potential clients have the possibility to contact Bridgis via a central email, phone number and visiting address. Bridgis also uses Twitter, LinkedIn, YouTube and Slideshare. The sources that are used for creating the product are mentioned

as is the intended use in combination with the most likely user type, user costs and references. A detailed manual is provided in the form of a video on YouTube (Bridgis, 2012).

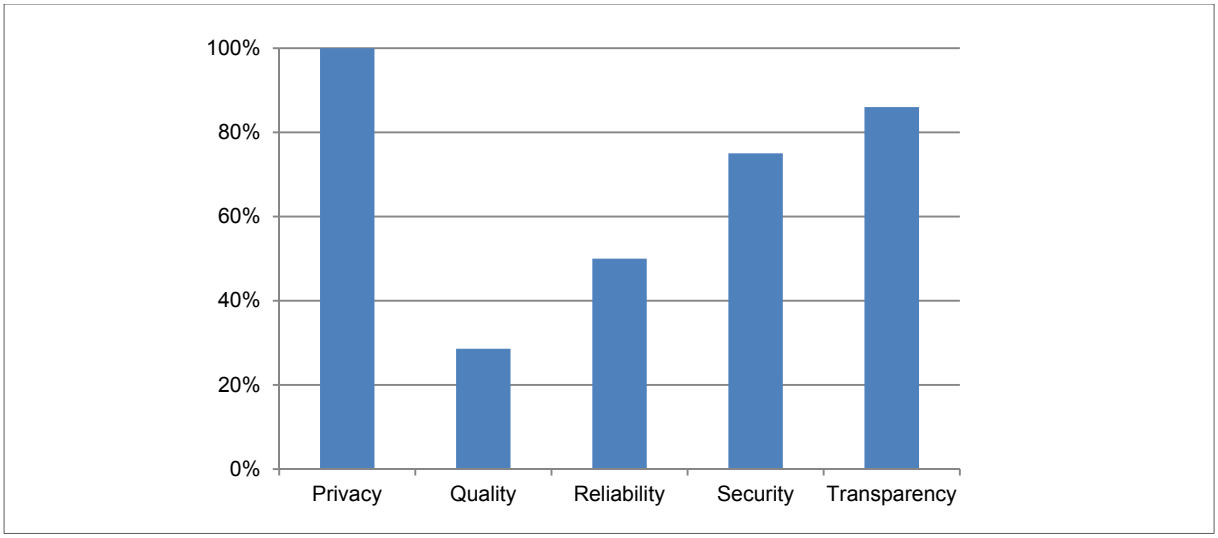
Because of commercial developments, Bridgis had rather not I would contact their clients and as there are no online user forums, the interview results have not been related to practise.
Bridgis did provide a report in which the developments of the use of the Populator over the years 2010, 2011 and 2012 are documented. Over the past two years, it shows a steady grow in the amount of individual users and a doubling of the number of user sessions. This is at the least an indication that the users are satisfied with the product.

4.3.2 MuRa by Nieuwland

Nieuwland describes itself as a results-oriented, knowledge-driven service organization that operates in different fields like training, re-integration, geo-information systems, green and environment (Nieuwland, 2009).

In the field of geo-information systems, MuRa (MuskRat) is an example of a product whose clients are dependent on it for their business. It is an application developed for Dutch Water Boards with which field staff can enter, mutate and retrieve location based data.
The primary use is to register information about muskrat catches and traps. Additionally it is used for example to exchange information about damages on location and the registration of employee timesheets.
A mobile device (smartphone) is used to enter or retrieve data by means of a predefined input structure and a map. Through an authorization scheme the user gets access to the application on the device.
For the use of the app and the service of Nieuwland an annual fee is charged, without limitations to the amount of users or data transfer. This is based on a fair use policy.
Currently the application is owned and used by six cooperating Water Boards and the service is fully managed by Nieuwland. Other Water Boards are users.

The interview with Nieuwland about MuRa was held on the twenty-sixth of June 2013, from 9.30 to 10.45 AM, at their office in Wageningen. Two persons, a service manager and a business manager, were present. The results are presented in graph 7.



Graph 7: Combined results after the interview with Nieuwland and the web research.

Personal data are not used in creating the product and as such MuRa fully complies to the Dutch personal data protection act.

Regarding data quality, Nieuwland, together with Grontmij, scores average when only looking at the commercial parties. The difference between the two is that the data suppliers for MuRa are also the data users. Apart from the background layers, the information that is used in the application is inserted and mutated by the clients themselves. Therefore, Nieuwland has little to no influence on the data quality. Because the data is not shared with other organizations, the lack of registering the data lineage and the metadata is unlikely to have harmful effects, all involved parties know where the data are coming from.

Nieuwland scores just above average on reliability. Information about the process is well documented and, according to policy, fully known by at least two persons. But in case of bankruptcy, the service would stop. Nieuwland has, however, made arrangements with their clients in which they send a closed envelop with the source code of the application every year, so the application can be rebuild.

In general, Nieuwland offers a good server uptime in their SLA (99,5%). During the development fase the client has agreed an uptime of 90%. Thus an incident in 2012 that took place during the weekend had no financial or other consequences.

When there is no connection, the field operator is still able to perform (most of) his duties. The application can make use of an offline map that is stored on the mobile device, together with basic information about the section the operator is working in. Data that is generated by the operator is stored on the device. When the device is online again, the data is send to the central server.

Nieuwland is one of the organizations that have experienced a minor security breach. There was no damage, but nevertheless they act on that experience and are planning to implement a formal security architecture framework, which will be completed in 2014. A second reason to implement a formal security framework is image. Clients like to know their data supplier takes security serious.

Safety risks have been identified and because security management at the moment is mainly externally oriented, Nieuwland believes the biggest potential threat is internal (employees), e.g. easy hackable passwords or a lost mobile device.

Both participants have taken ample time to inform me about the application and more. Nieuwland has two contact persons for MuRa, who can be contacted by phone. Also central contact information (mail, phone, address) is given. There are not that many data sources Nieuwland needs to incorporate in the product, but they can be found on the website, together with the intended uses and users. Potential clients need to contact Nieuwland to get information about the user costs and only a general description about the working of the application is given online. No direct references are placed on the site, but visitors are redirected to two articles covering MuRa (Nieuwland, 2013).

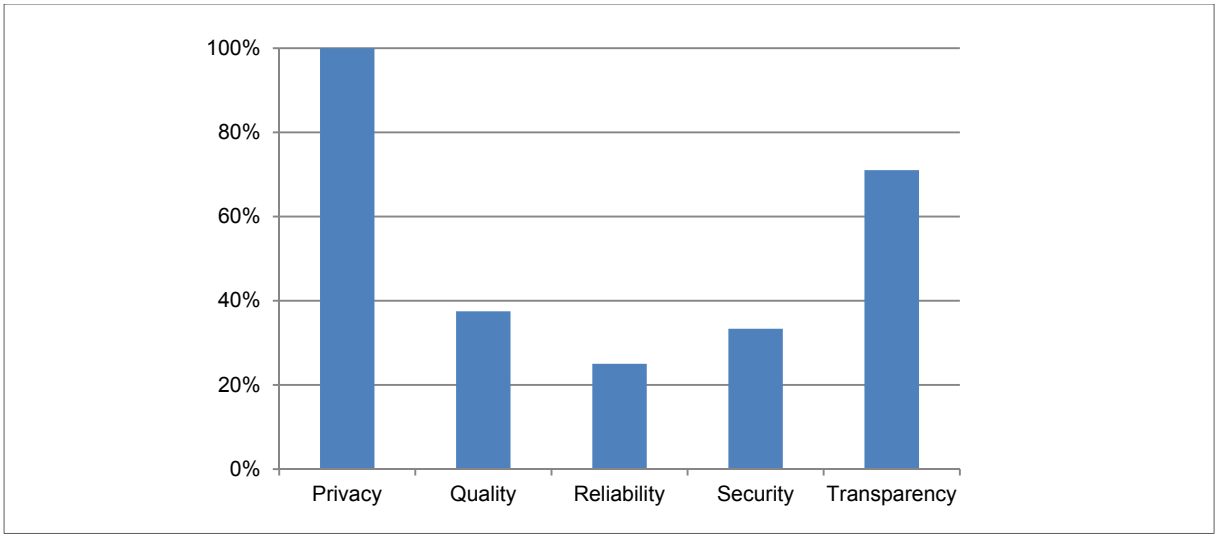
A new user of MuRa was contacted to comment on the product. They are still testing the product, but the results so far are promising. They see it offers benefits to the field staff in that they administrate their catches and means more accurate, as well as to management who can analyse the information faster and more accurate.
They said that inserting objects in the map when offline can show a deviation compared to working online. Also working with a mobile device takes some time getting used to for the field staff. Nieuwland is putting great effort in helping the people get accustomed to working with the application.

4.3.3 Buienradar by RTL Nederland

The RTL Group is a media company that owns commercial television and radio stations in Europe. A few years ago it bought Buienradar to facilitate its own weather forecast on television and radio.

Buienradar provides graphical information about rainfall on a specific location and a specific time. The information can help clients organize their daily outdoor activities (e.g. picnics, farming or traveling). Before sending the imagery to the users, the raw data is processed, taking in account several aspects that influence rainfall on ground level, like air humidity and temperature.
There are no costs linked to the use of Buienradar. Revenues are primarily generated from advertising. In the future Buienradar means to provide even more accurate information by incorporating data from neighbouring countries.
Users are not registered so it is impossible to precisely tell who uses the product. The known users are very diverse and range from civilians and farmers to police and governments.

The interview with RTL Nederland about Buienradar was held on the twelfth of June 2013, from 9.30 to 10.45 AM, at their office in Hilversum. The product manager of Buienradar was present. The individual results of the interview are shown in graph 8.



Graph 8: Combined results after the interview with RTL Nederland and the web research.

Also in Buienradar, no personal data have been used and so the product complies to the privacy guidelines.

RTL Nederland scores best on quality compared to the other commercial participants in this study. The quality is checked continuously by an operator. But because the procedure changes frequently, depending on different factors related to weather condition, the procedure is not documented and knowledge is transmitted verbally and gained on the job. For Buienradar, correctness and temporal accuracy of the information are essential but data quality is not guaranteed. The data model is complete, but nothing about data lineage or metadata is documented.

Although the production process is known by more than one person, RTL Nederland scores very low on reliability. In case RTL Nederland would stop operating, no arrangements have been taken to continue the service. The production process is documented, but specialist knowledge is needed to create usable information. This means the service would stop completely without one of the current operators sharing his knowledge. Also, no guarantees for server uptime are given and server uptime is not measured.

Also on security, RTL Nederland scores low. No formal security architecture framework is implemented or planned to be implemented and there is no active security management. Because the data is free for non-commercial use and only one way interaction, according to RTL Nederland security risks are low. The risks have been audited and the results are taken in consideration in the development of the product. Hacking and viruses are seen as the biggest risks.

The main reason reliability and security are low on the priority list is that RTL Nederland has not experienced difficulties in these fields and thus the urge to act is low. If Buienradar would be offline, users would experience little safety or financial consequences. But the product has by far the most daily users of all commercial products that participate in this research and delivery problems would affect a big number of people in their daily lives. RTL Nederland is aware of this and is formalising the production process of Buienradar.

The participant was open in answering the questions and more than ready to give information about the product. Contact information on the website is limited to a central email address. To stay in contact

with users, Twitter and Facebook are used. Information sources are mentioned on the website. Intended use, user costs and a manual on how to use Buienradar are absent. Not that it is being missed, its purpose is obvious, the use is free of charge (for non-commercial use) and a manual is not needed, using the service is very intuitive (Buienradar, 2013).

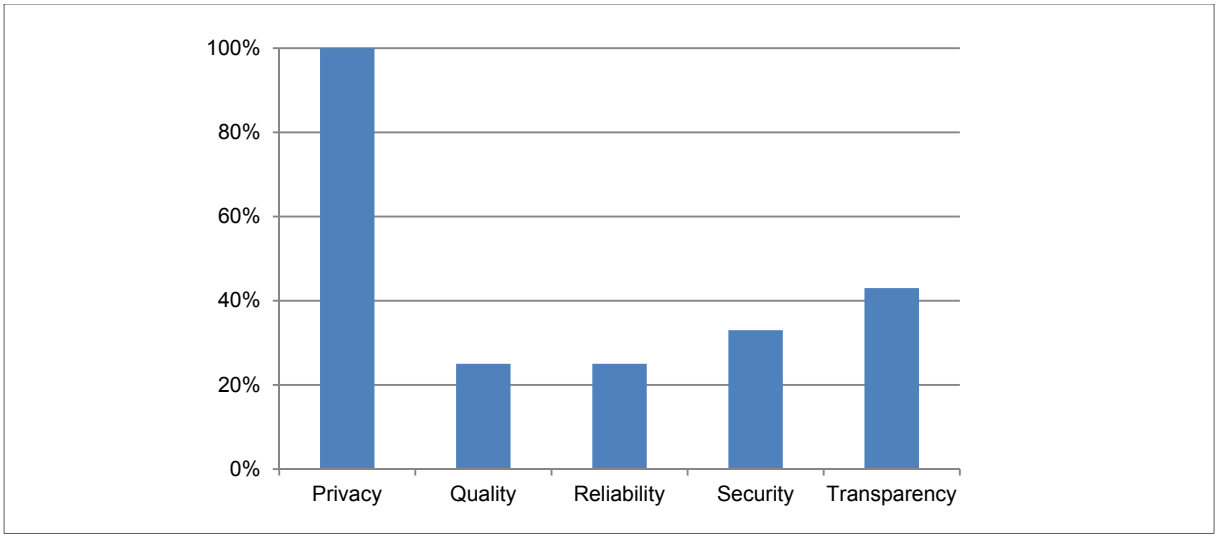
Buienradar has a low average score on quality, reliability and security. Most users will not even think about security of the service, but they are interested in quality and reliability. On the internet, a lot is written about Buienradar. Comments are mainly positive, especially about the accuracy of the rainfall prediction (quality of the information). Users are less satisfied with the reliability of the service and complain about crashes, errors and slowness (e.g. Google, 2013).
Weather, and rain in particular, has been and will always be a topic people will want to be informed about, because it affects every day life in a very direct way. And it is no surprise that Buienradar faces increasing competition. Although it is the best known service, it will have to deliver better quality of service and data than other, similar applications (e.g. Buienalarm, Rainy days).

4.3.4 Radar by Grontmij

"Grontmij provides consultancy, design & engineering and management services in a broad range of market sectors related to the built and natural environment" (Grontmij, 2013).

Radar is developed to help consultancy and governmental agencies get easy access to information that is used in studies prior to creating development plans. The goal of Radar is making the process of spatial planning more cost and time efficient.
Radar combines different data sources, ranging from environmental data to data about hazardous objects. The user (graphically) indicates an area of interest and as a result, a report is delivered with information about the specified area that is relevant to spatial planning. Also in the report all used sources are listed, together with an abstract of the metadata, so the user has insight in i.a. the source and the age of the used data.
For the use of Radar an annual fee is charged.
The product is less successful than expected and at this point the number of clients is limited. Grontmij is looking for ways to increase the amount of customers and is in the process of redeveloping Radar. One of the ideas is to connect Radar to Geoweb, a web viewer application. This enables customers to integrate their own data in the analysis. Also Grontmij wants to add other layers and functionality to broaden the uses of Radar.

The interview with Grontmij about Radar was held on the twenty-eighth of June 2013, from 9.00 to 10.00 AM, at their office in De Bilt. The project leader of Radar and one intern were present. The combined results are shown in graph 9.



Graph 9: Combined results after the interview with Grontmij and the web research.

Radar has the lowest overall score of all products. On privacy the score is perfect, but, as with most NI in this research, Radar does not make use of personal information.

Quality is being checked irregularly and manually without a clearly documented procedure. Because of the limited number of clients, data updates are irregular and only performed when needed. The product manager at Grontmij has the ultimate responsibility on the quality of the product. At this point, no regular reports on quality are being written nor have key performance indicators been determined. The product team only documents the part of data lineage and metadata that is part of the resulting report that is send to the client. The data model is completely documented, but will most likely change significantly in the near future because of the redevelopment of the product.

The score on reliability is also low. Although the process of creating the product is well secured, the other elements of reliability are not.

Individual parts of the process are known by different employees and the whole process is well documented. Also, new employees are being trained by experienced colleagues. Because of the limited number of clients, no arrangements have been made to secure the service. So, in case Grontmij would go bankrupt, the service would stop. Also, server up or down time are not being monitored.

On security, Grontmij has not implemented a formal security architecture framework and no security management is active. The safety risks for Radar have been identified and are acted upon during the development phase. Safety measures that have been taken are firewalls and antivirus programmes and regular software updates. Clients need to login and are not able to connect directly to the data server.

Grontmij scored insufficient on transparency. Although the interview was very informative, the website of Grontmij on Radar is not giving the basic information a (potential) client needs. It is possible to contact Grontmij via the central address but on the website, an incorrect contact person is given. Also no used sources, no user costs and no references of satisfied customers are given. However, the intended use and a general description about how the application works and what the user has to do to get results are given.

The product is being revised and repositioned, and with that, information on the website, client contact and other communication and sales activities will get extra attention.

Because Grontmij is redeveloping and repositioning Radar, from the Grontmij perspective, it is not desirable to contact clients about their experiences with Radar.

4.3.5 BRT by PDOK and Kadaster

PDOK (public service on map) enables users to access digital geographical data via official PDOK webservices. More than thirty web services (aimed at digital mapping) are available to the general public, private companies, organizations and the public sector (PDOK, 2013a). PDOK provides the service, while the distributed data is created and owned by various governmental agencies.

Kadaster (the Dutch Cadastre) registers property and geographical information and is the owner of the BRT (Basic Registration Topography) (Kadaster, 2013). BRT is one of the datasets that are distributed through the webservices of PDOK.

PDOK processes five to seven million map requests a month and with an average of 450.000 requests a month, the BRT is in the top five of requested products (PDOK, 2013b). BRT is one of the two DaaS that are part of this research. It is a topological product that is mainly used as a reference layer, but can also be used for e.g. planning forest maintenance by the Dutch state forestry. All government organizations are required by law to use available data from BRT to exchange information based on a geographic base (Woudenberg, 2012 and Overheid.nl, 2013a).

The use of BRT through PDOK is free of charge and it is allowed to copy and redistribute the BRT information, also for commercial purposes. When reusing or redistributing the information, a reference to Kadaster should be made.

PDOK has three service levels: Basic, Educational and Fair Use (figure 11). Basic is intended for governmental organizations. They get the highest level of service because they are required to use data that is being distributed by PDOK. Educational purposes have medium service and ‘other’ users are entitled to limited service. Looking at the registered users, governmental agencies are by far the biggest users of PDOK.

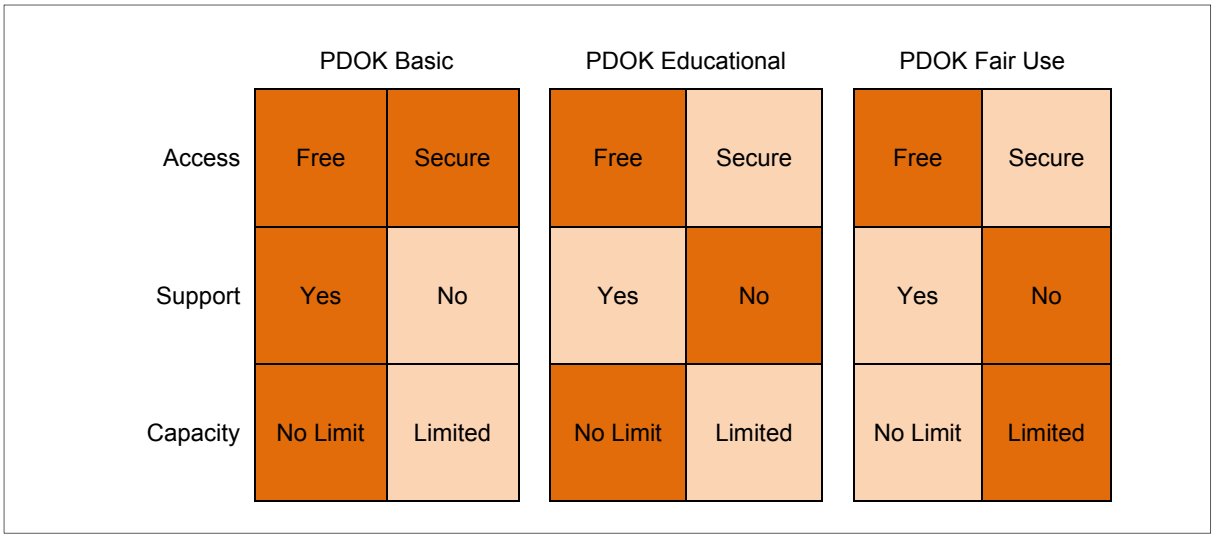
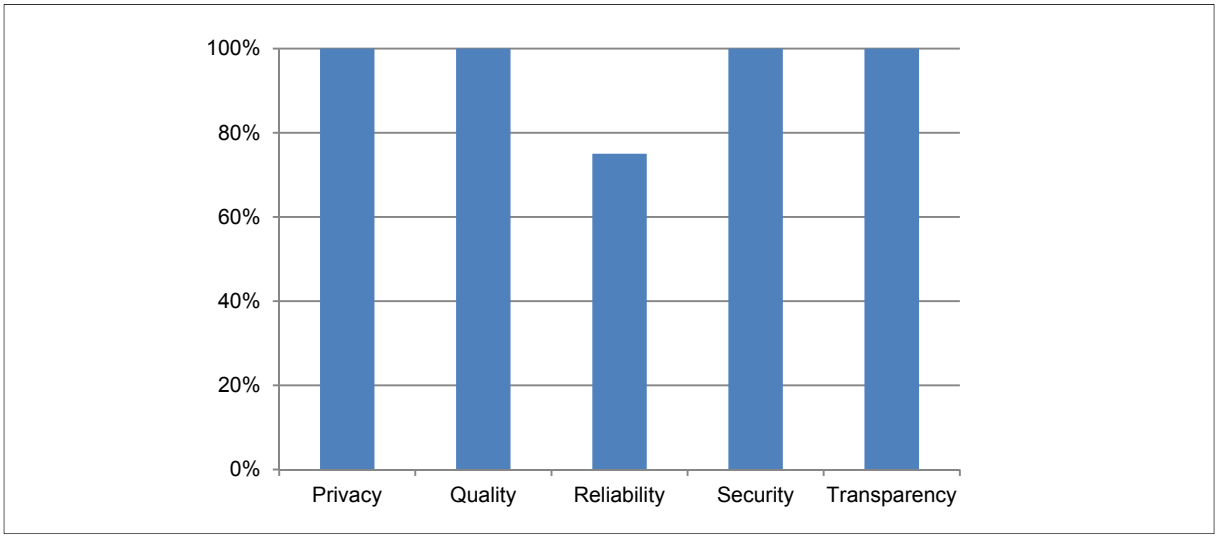


Figure 11: Matrix that shows the level of service PDOK offers for each category (based on: PDOK, 2013a).

The interview with PDOK and Kadaster about BRT was held on the nineteenth of June 2013, from 14.15 to 15.45 AM, at their office in Apeldoorn. The product manager of PDOK handled the questions about security and reliability of the PDOK services and the senior advisor of Kadaster answered questions about privacy and quality for the BRT. The results are presented in graph 10.



Graph 10: Combined results after the interview with PDOK and Kadaster and the web research.

The BRT service scores best of all products, even 100% on four of the five aspects of usability. Kadaster does not use private information in the process of creating BRT.
Quality of the BRT data is checked on a regular basis. Checks are performed partly automated and partly manual and are fully documented and published through the Kadaster website. The process and the documentation are audited yearly. Once every three years an external organization checks the data quality. The report of that check is published on the Kadaster website. Quality management focusses on correctness, completeness and temporal accuracy and the director is informed quarterly

about the progress. The data model is fully documented and published on the Kadaster website. Also data lineage and metadata are fully documented and can be requested for.
PDOK scored very well on reliability. The only element that is not covered is the continuity of the service in case the organization would cease to exist. But, it being an governmental agency, that would be very unlikely.
PDOK uses a formal security architecture framework, security management focusses on both internal and external threads and the director has overall responsibility.
The participants from PDOK and Kadaster were very well prepared for the interview and they took plenty of time to inform me about the service and the product. On the websites of PDOK and Kadaster, central contact information is given. Kadaster also uses Facebook, Twitter and YouTube to inform users and a BRT discussion group is created on LinkedIn. Because BRT is a base product, used to create other products, no sources are described as part of the production process. PDOK and Kadatser give examples on uses of BRT and PDOK provides complete manuals about how to use their services. No costs are calculated for the use of BRT and on the site of PDOK, different clients give information about their experience with their services.

Kadaster (data owner) has given insight in an unpublished survey report in which the results of a customer satisfaction survey, also on BRT, are published. A number of aspects that were part of the investigation of Kadaster match the aspects used in this research to measure usability and deal with quality, reliability and transparency. Other aspects the Kadaster looked at were e.g. ease of use and how the data connects to the processes and software of the client. Three kinds of users took part in the investigation: governmental agencies, real-estate and 'other' commercial parties.
BRT scores 7.1 overall, where governmental agencies gave both the highest (tax administration) and the lowest scores (provinces). The overall score indicates that the data is indeed 'very usable'.
Also on the site of PDOK (service provider), user stories about their experiences with the PDOK services are posted (PDOK, n.d.). Clients are enthusiastic about the ease and the availability of the services.
Also, the site of PDOK has a three star 'drempelvrij.nl' label, which means the site is easily accessible and of good quality (Stichting Waarmerk drempelvrij.nl, 2012). Although It does not say anything about the quality of the data service itself, it does indicate that PDOK is putting effort into quality of service.

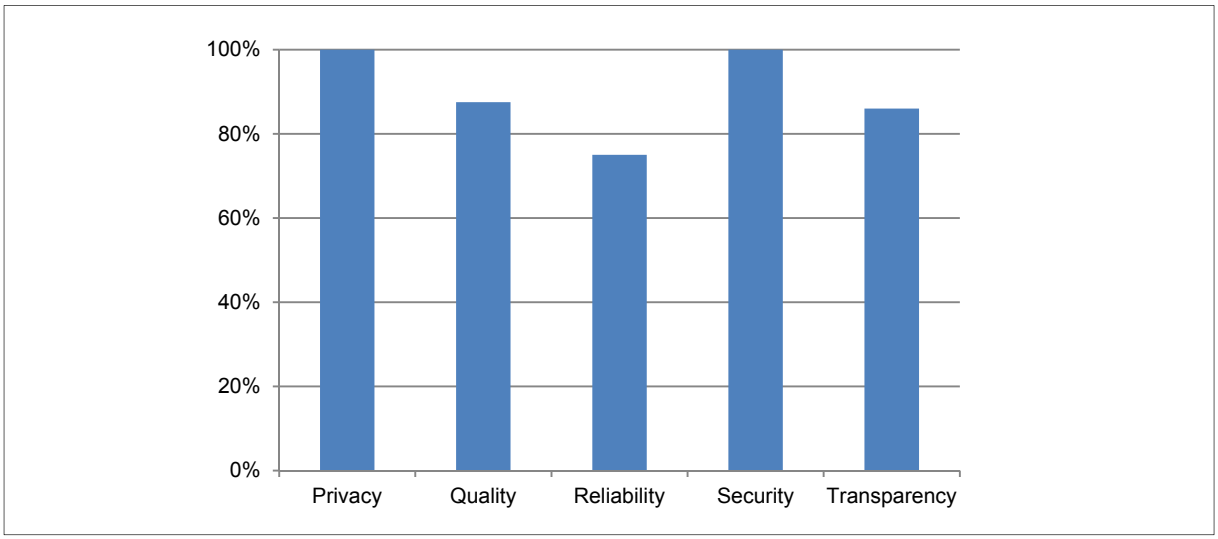
4.3.6 NWB by RWS

RWS (Rijkswaterstaat) "is part of the Dutch Ministry of Infrastructure and the Environment and responsible for the design, construction, management and maintenance of the main infrastructure facilities in the Netherlands" (RWS, 2013).

NWB (Nationaal Wegen Bestand) is distributed as a DaaS. It is a geographic database of all roads in the Netherlands that are managed by the central government, provinces, municipalities and water boards. It is used for i.a. asset management, navigating and to link traffic accidents.

The three top users are RWS itself, police control room and local authorities. The (commercial) use of the product is for free.

The interview with RWS about NWB was held on the first of July 2013, from 9.40 to 11.00 AM, at their office in Maastricht. The product leader of NWB was present. The usability scores based on the interview are presented in graph 11.



Graph 11: Combined results after the interview with RWS and the web research.

Like all government agencies, NWB has a good overall score. In the production process of NWB, RWS does not use private information.

Quality management focusses on correctness, completeness and temporal accuracy, but is only ensured on the level of team manager. Also results are not reported in writing, which has a negative effect on retrievability of quality development. The data model, data lineage and metadata are fully documented and the Rijkswaterstaat Centrale Informatievoorziening (CIV) is responsible for the documentation. The data model, and thus the content of NWB, has changed little over the past eight years. Instead, other sources have been linked to it.

The entire production process is known by multiple people and completely documented and the guaranteed and achieved server uptime are very good. However, no arrangements have been made concerning the continuity of the product in case RWS would cease to exist, although this is very unlikely.

NWB light is an aggregated version of NWB. It is distributed by the Rijkswaterstaat Centrale Informatievoorziening (CIV), which is responsible for the development and availability of information within the RWS organization. The NWB product is distributed by PDOK, the same organization that distributes the BRT dataset. Therefore, NWB received the same score on distribution related questions as BRT.

Apart from PDOK, RWS has an active security management, aimed at both external and internal threads. Data are stored on a server that cannot be accessed by unauthorized persons and regular security sessions are organised. The director has overall responsibility and is informed about progress

after each session. Infiltration and unauthorised access of RWS data sources are the biggest security concern.
RWS scores good on transparency. The RWS website provides central contact information together with information about the intended use of NWB. Both WS and PDOK provide complete manuals on how to use the information. No costs are charged for the use of NWB and no references are posted on the site.

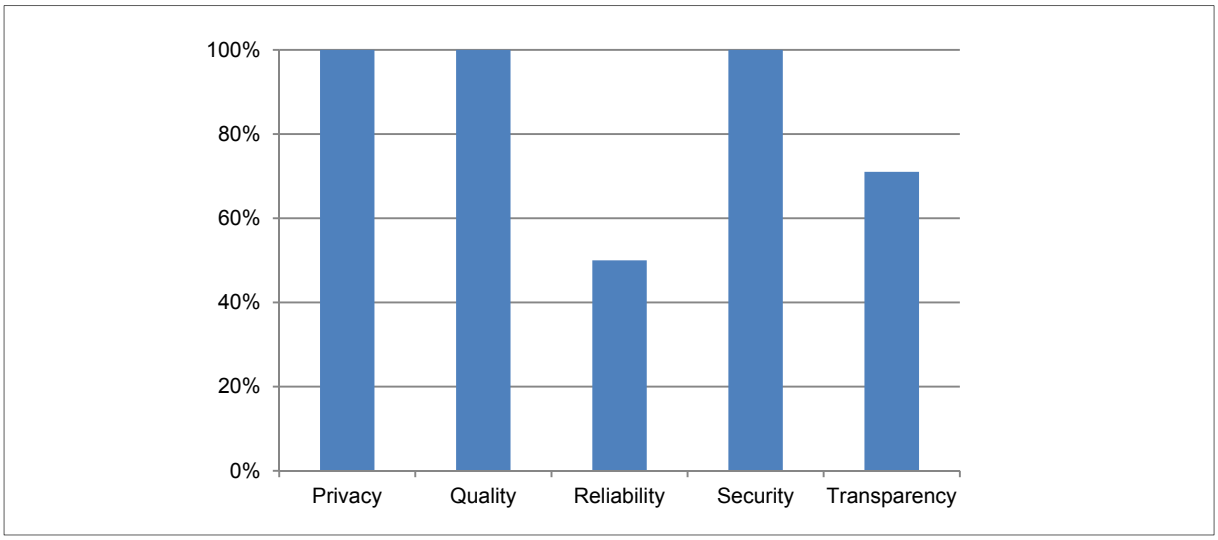
RWS has initiated a research to customer satisfaction on NWB. The research has not yet been finished. A governmental user did confirm the positive results of the questionnaire by indicating the data is adequate for use and RWS is easy to reach in case of questions.

4.3.7 Risicokaart by GBO provincies

GBO provincies (Gemeenschappelijke Beheer Organisatie provincies) manages common ICT-information services for the Dutch provinces and aims to contribute to the development of these information services (GBO provincies, 2013)

The initiative in creating Risicokaart was taken after a fireworks explosion in Enschede in 2000 and its main purpose originally was to inform civilians about safety risks in their direct vicinity. Nowadays Riskmap also facilitates professional users (e.g. emergency services) by recording objects that are not intended for the general public (e.g. major gas lines). The main users are civilians, followed by companies and municipalities, who use it in the process of spatial planning. Usage is not charged. GBO provincies use information that is provided by local governments in creating Risicokaart.

The interview with GBO provincies about Risicokaart was held on the fifteenth of July 2013, from 10.00 to 11.30 AM, at their office in The Hague. The coordinator data management of Risicokaart was present. The results are presented in graph 12.



Graph 12: Combined results after the interview with GBO provincies and the web research.

Risicokaart scores very good compared to all other organizations, high on privacy, quality and security and average on reliability and transparency. Despite the good results, GBO provincies is still working on improvement in the field of quality.
GBO provincies does not use private information in the production process of Risicokaart.
Quality management focusses on correctness, completeness and temporal accuracy. Findings are reported monthly (internal) and quarterly (external). Momentarily, quality checks are performed irregularly, and are only partly automated. The procedures are not documented. GBO provincies is putting effort in developing a further automated quality check and the complete documentation of quality related procedures.
The data model is fully documented, but in the eight years it exists it has only grown. GBO provincies is going to check if all individual parts of the data model are still valid. For this, the relevant laws and regulations and the different partners are consulted.
The score on reliability was average. The full production process is known by more than one person and documented, on functional and technical level. No arrangements have been made should GBO provinces cease to exist, but suppliers are required to ensure continuity of their products. Also both data and service are backed up off site. The best guaranteed and achieved server uptime was only 98%.
GBO provincies use a formal security architecture framework and security management is focussed externally and internally. Off all participants of this research, GBO provincies is the only organization to organize an awareness training for their employees and to appoint a security coordinator. The risks have been identified and have been made part of their annual plans.
On the website of GBO provincies, two contact persons are given. They can be contacted by a central mail address. On the site of Risicokaart, only the (same) central email address is given. On the website, the organizations that deliver information for Risicokaart are mentioned and the intended uses, with the intended users, are clearly defined. The use of Risicokaart is free of charge. No references are put on the site. Risicokaart is so easy to work with that no manual is needed. The information that Risicokaart provides is described extensively, so the user knows how to interpret the result.

A provincial coordinator was contacted to provide feedback on the results of the interview. Amongst others, the coordinator is responsible for the communication between the users and the GBO provincies organization. They collect user problems and pass them on to the GBO provincies organization. The term 'user' identifies two groups, namely the organizations that supply the application with data about potential risk objects and the people or organizations that use Risicokaart to locate potential risk objects. The coordinator confirmed the need to redesign the input application and check the data model, because the users who provide the data, find the application difficult to use and the data model inconsistent. Information users are pleased with the application, both on quality and availability.

CHAPTER 5
Conclusion and discussion

5.1 Summary

5.1.1 Introduction

Longley, Goodchild, Maguire and Rhind (2005, p.12) mention that a "… characteristic of information is that it is easy to add value to it through processing, and through merger with other information". And that is exactly what companies are doing. They use the available (online) data to create a product that they can sell and make profit of: information.

By combining and processing the available online data with or without data from other sources, companies can meet a market demand and make the produced information available as a service. The product is Networked Information (NI). It is generated real-time from at least one source available via the internet and it is distributed online. It is a result of a Model as a Service (MaaS), which consists of the elements data, service and application.
Dwivedi and Kulkarni (2008, p.615) describe the process of creating NI, which "entails capturing raw information across sources, its storage and validation and providing analytical reports to the clients". The process has three actors. The supplier of the data used to generate the NI, the producer of the NI and the client, who buys and uses the NI.

The benefits of NI are much like the benefits for DaaS. First, it is cost effective, you buy what you need. Second, it increases flexibility because accessing the data is location independent. And third, "access to the data is controlled through the data services, which tends to improve data quality" (Rajesh, Swapna and Shylender Reddy, 2012, p.28).
According to Rajesh, Swapna and Shylender Reddy (2012, p.28) the drawbacks are that the client is dependent "on the service provider's ability to avoid server downtime" and that "the consumer is really just "renting" the data, … generally the data is not available for download".

Some NI products are used daily by large amounts of people. These products have a great social impact just because of the amount of users. Other products have a higher safety or economic impact because of use in healthcare or disaster management (saving lives) or because companies are depending on the information for their business (making money). This means information should be usable for a specific purpose (e.g. making money).

5.1.2 Problem statement and research questions

The phenomenon NI is emerging and its use is growing rapidly, privately as well as commercially and in government (e.g. risk management). This means that the social, economic and safety dependency on NI products is growing. That's why it is important to know what makes the information fit for use.

Ultimately, if the usability of the NI product can be measured, and the aspects that influence it are known, than producers of NI will be motivated to improve the usability of their product.

Because NI is a service and it is about information use, the aspects involved are related to (geo) information in general and to service oriented architecture (SOA), specifically to data, information and model as a service, for the distributed and service parts.

The main question for this research is:
Which aspects of networked information influence its usability and can the aspects be used to measure and compare the usability of networked information products?

The main question is divided in four Sub-questions:

5. What is usability in relation to NI?
6. What requirements does NI need to have to be fit for use?
7. How can the requirements for the fitness-for-purpose of NI be measured?
8. To what extent can the producer of NI positively influence these aspects?

5.1.3 Literature review

"Usability is the degree to which something - software, hardware or anything else - is easy to use and a good fit for the people who use it" (User Experience Professionals' Association, 2012). Designing a usable product starts with usability engineering, which, according to User Experience Professionals' Association (2012) is "the disciplined application of usability practices to assess the needs and abilities of users, in conjunction with the business requirements, practices, and processes of an organization". In other words, defining the usability of a product starts with finding out what the users' needs are.

Because of the great quantity of digital available personal related data from social media and government and companies like Google and Microsoft who collect and use personal data in their business, privacy is a big issue in NI. Also data needs to be true (correct), complete and up to date to be used in decision making. These three elements come together under the aspect quality of a data set.
Delphix (2011, p.3) mentions both timely and secure as required characteristics of DaaS. Clients should be able to access their data when they need it. In this context, timely is part of reliability, in that being able to access the data when needed means that the data should always be accessible. "The highest quality information is useless if it cannot be seen so that it can be acted upon" (Tortorella and Driscoll, 2005, p.3). The continuity of the data supply in case bankruptcy of the producer or a data supplier and availability of information by avoiding server downtime (Rajesh, Swapna and Shylender Reddy, 2012, p.28) are also part of reliability. Because information is being produced and kept external, by an outside party, clients must be able to rely on their sensitive data (e.g. business-critical information) being secure and unavailable for any unauthorized access and use.

For being able to assess the above aspects, the producer needs to display a certain amount of transparency. If clients are not given access to information relevant for measurement, they cannot assess the usability of the product.
With this, five aspects have been identified that influence the usability of NI; **P**rivacy, **Q**uality, **R**eliability, **S**ecurity and **T**ransparency.

5.1.4 Methodology

This research is about information as a product and the degree to which it can be used for a specific purpose, based on facts and measurement instead of emotion or a deeper meaning. Therefor the research method is quantitative. Data about the NI products is accumulated through a combination of web research and a questionnaire (appendix A). The products are compared using a benchmark.

The questionnaire is divided in three parts and consists of 27 questions and a number of follow-up questions. In the first part the participant is instructed and terms used in the questionnaire are defined. Part two contains easy and neutral product descriptive questions. The questions in part three deal with the aspects that influence the usability of the NI (chapter 2). This part consists of closed questions, with categorical response choices, and open follow-up questions.
The highly structured level of the closed questions make the results very suitable for quantitative analysis (Gray, 2004, p.215) in a benchmark. The function of the open questions is to verify the given answers as far as possible and to allow the participant to clarify his answers.

What information about a NI product is publicly available is determined through web research. The amount of information available determines part of the final score of the aspect ‘transparency’.

The subject (the population) of this study is networked information (NI), which is generated real-time from at least one source available via the internet and it is distributed online. The sample used in this study is randomly selected from the population, with the exception that no two selected products serve the same purpose.
The sample consists of five NI products and two data as a service (DaaS), which can be used as an input for NI. The DaaS are not part of the population, but share many of the NI properties. They are included to see if the results show significant difference in comparison to NI.

5.1.5 Research limitations and delimitations

This thesis focusses on NI products that have a geographical component. The main reason for this is that the use of information with a geographical component is so common that most people do not even realize that they are. “Almost all human activities and decisions involve a geographic component” (Longley, et al., 2005, p.8).

For this research, the accuracy of the information that the NI producers deliver is vital. It is assumed that the participants (producers of NI) will truthfully answer the survey questions. But there may well be

a limitation regarding business sensitive information (e.g. server downtime). Companies might not be inclined to share this kind of information.

The sample used in this study is randomly selected from the population, with the exception that no two selected products serve the same purpose. The hope is that participants (being the producers of the NI) are more inclined to provide information when their direct competitors do not participate in this research.

Benchmarking is a tool for evaluation and improvement. Usually, products are compared to other products which are “recognised as the best within the area” (Andersen and Pettersen, 1996, p.3). But, because the method of benchmarking in this study is used purely as a manner to implement the theory of measuring the usability of a NI product according to the described aspects, instead of looking for and comparing to the best product, the NI products will be compared to the baseline that is defined as a result of the literature study.

The subject is the usability of the networked information (e.g. quality, legal issues and reliability) and not the software, and therefore user experience (e.g. layout or satisfaction of user requirements) is not part of the research.

5.1.6 Findings

The interviews were conducted within a few weeks’ time. All participants were very cooperating most participants were well prepared. The interview started with a short introduction about the research, the time necessary for the interview and what would happen with the results.
After all interviews were held, the individual results, were send to each of the participants, to give the opportunity to correct any inconsistencies before the results are published.
For only three of the seven products, users have been contacted personally to comment on their experiences. In case of one product, the internet was used to find user experiences, in another case, an internal customer satisfaction survey was used. Two organizations had rather not I contacted their clients because of business developments. This is respected and for these products, users have not been contacted.

The gathered data was used to compare the seven products (five NI and two DaaS) using a benchmark. The products have been compared to each other and to a theoretical standard (baseline) instead of the product recognised as the best within the area, as is usual with a benchmark.
The baseline that is used (appendix B) is a working example. For every user and every use the focus and therewith the values for the different aspects, can be different.

On the basis that in this study all questions weigh the same, and compared to the working example baseline, the service (DaaS) of PDOK in combination with the data of Kadaster has the highest score

(twenty-five out of twenty-six) of all products. Looking solely at the NI products, Risicokaart of GBO provincies (twenty-two out of twenty-six) would be the most usable product.

It is striking that the products of the government agencies have a better score than the companies have. Especially in the field of data quality and security, government agencies stand out.

5.2 Conclusion

The conclusions of this research are divided into two groups, namely conclusions based on the findings of the interviews and the benchmark, and secondly answers to the research questions. They are presented respectively in this paragraph.

5.2.1 Conclusion concerning the interview

Based on the results of the interview (chapter 4), two conclusions can be drawn. First, the level of usability of the majority of NI is medium to high. Second, information products made and distributed by, or on behalf of government agencies are more usable than the NI of the companies, especially in the field of data quality and security.
Although unleashing the full power of statistics on the results is meaningless, because of the small size of the sample, these conclusions seem to be valid for information products in the Netherlands and possibly international.

In chapter 4, a small set of characteristics has been identified for each of the products (table 1). The set consisted of type of product, the intended use, number of users, type of producing organization and what kind of influence the product has (following the use of the product). Of these, only the type of organization, government or company, seems to be of influence on the usability of NI products. According to the total scores (graph 5), companies score about 50% to 60%, where governmental organizations score 85% or higher. Especially data quality and security of governmental products is better arranged than products produced by companies. Products that do not fit the needs of users, will face the concept of market mechanism, survival of the fittest. The product will either stop being produced, because it is not profitable, or it will be improved.

But, does the image of government vs companies correspond to reality and, if yes, how could this difference be explained? First, it is good to indicate that the approach chosen in this research is pure theoretical. The aspects have been determined as a result of a literature study and the information about the products has been collected from the producers of the NI. The users of the NI have barely been involved in the research. This could affect the outcome.

The first motivation could be that government agencies are ahead of companies in incorporating standards. Although the government is generally not known for being leading when it comes to

(information) technology, it can also be a drive for change. The Dutch government committed itself to the Infrastructure for Spatial Information in Europe (Inspire) directive, which “aims to create a European Union (EU) spatial data infrastructure” (European commission, 2013). PDOK, the data distributor of the Dutch government, works according the Inspire directive and so the service meets the directives requirements. True, this does not influence the quality of the products itself, the production of for example BRT or NWB did not change because of Inspire, but topics like metadata, availability and standardization, which impact both quality and reliability, are incorporated in the Inspire directive.

More important however, are the motives of an organization for producing NI, set against the time and money that has to be invested. “Good data cost money” and “considerable time and effort is required” (Long, Roberge, Lamicela and Murugesan, 2004, p.7). Every producer, government and companies alike, has to “earn” their investment back in one way or another. The goal can be different. Companies try to make profit, so the time and money invested should resolve in more money. Because of this, good will often be good enough. This applies to other aspects of usability as well, not only quality and it is only natural for companies to deliver products that meet the requirements, but are not (much) better than the customer asks (and pays) for. During the interviews, most respondents indicated that they are improving one or more aspects that influence NI usability, or at least are planning to do so. Reasons that were given, were often aimed at (future) clients. Government agencies on the other hand, are less money driven. Although finance is getting ever more important in the operating of governmental agencies, the main goal is not to make money by selling products, that is done through taxation. A government should have a higher goal, or at least it has a greater responsibility towards its citizens.
The Dutch government released a memo in which the national ‘iStrategy’ is described (Rauch, 2012, p.1). The goal is to create “a leaner, more efficient Central Government through Information and Communications Technology (ICT) and Information Management (IM)”. Amongst other items, specific attention is paid to data quality, privacy, security and reliability. These aspects are described as important in this strategy because of the reputation of the government and its functioning. Citizens should “feel they can trust Central Government to store and use any data relating to them with the greatest of care, making sure that it is accurate and used in accordance with the law” (Rauch, 2012, p.8) and at the same time, the government needs a solid IS to be able to meet the intended outcome, which is “increased efficiency and reduced costs” (Rauch, 2012, p.1).
Although the memo looks sound and well thought through, money and time are limited in government as well. A lot depends on the quality of management and the individual employee.

Concerning transparency, companies have commercial interest in transparency, to attract as much clients as possible. The government on the other hand is required by law (wet openbaarheid van bestuur) to inform (Overheid.nl, 2013b). In this case, the motivation is different, but the outcome is the same.

Of the seven products that were compared, two are DaaS instead of NI. In chapter 3.3.2 I indicated that the DaaS are included to see if a significant difference with NI is shown in the results. The DaaS (NWB and BRT) do score very good, better than most NI, but this research did not produce enough information, from which definite conclusions can be drawn.

5.2.2 Conclusion concerning research questions

In chapter 1.2, this thesis started with laying out the main and sub-questions. With the literature review and the interviews finished, in this chapter the answers to the research questions are displayed in bold printing after each question.

Main question: Which aspects of networked information influence its usability and can the aspects be used to measure and compare the usability of networked information products?
Five main aspects are of importance for the usability of NI. These are; Privacy, Quality, Reliability, Security and Transparency. And yes, these aspects can be used to measure and compare the usability of NI products.

This answer is explained in more detail with discussing the four sub-questions below.

Sub-question one: What is usability in relation to NI?
As I have described in chapter 2.1, “Usability is the degree to which something - software, hardware or anything else - is easy to use and a good fit for the people who use it” (User Experience Professionals' Association, 2012). In other words, defining the usability of a product starts with finding out what the users’ needs are.
The essence of NI is that it is a service which provides geographical information, this means that it has the combined characteristics of Geographical Information in general and service oriented architectures, specific data and information as a service. Combining this with the primary use of NI, which is to assist in decision making, results in the basic users’ needs. **NI has to be legal and of good quality and the service needs to be secure and reliable. On top of that, to be able to assess the above, the producer needs to be willing to provide the user with inside information.** This translates to the five aspects **P**rivacy, **Q**uality, **R**eliability, **S**ecurity and **T**ransparency.

Sub-question two: What requirements does NI need to have to be fit for use?
Based on the five aspects, the requirements of NI can be defined. But as there are dozens of different NI and many more individual users, all with their own specific needs and wishes, it is not possible, or better yet, desirable, to define one set of general requirements that apply to all NI used in all applications.
It is basically up to the user to specify its needs. For example, although a NI should always comply to privacy legislation, in some cases privacy is not an issue, because no personal data are used in creating the product. In other cases, a NI does not need an elaborate security architecture. The requirements should be tailored to the situation.

On the other hand, assuming that the usability will be assessed on the basis of the method used in this research, it is possible to define a basic set of questions, from which a number of questions that apply to that specific situation, can be selected. The requirements, in the form of a baseline like in appendix B, are to be defined by the user.

Sub-question three: How can the requirements for the fitness-for-purpose of NI be measured?
As is described in chapter 3, in this research, measuring the usability of NI was done by interviewing the producers. The users of NI have virtually been ignored.

This approach has advantages and disadvantages. Firstly, the closed questions make it easy to score and compare the different products. At the same time, closed questions leave no room for subtleties.
Secondly, there is the question of who to contact for information, producers, users or both? Approaching the producers should, in theory, produce the most accurate information. A producer has all knowledge about the production process and the way the produced NI is made accessible. On the other hand, producers also have business motives and it might not be in their interest to disclose all information, for instance about server downtime or data quality, because this might damage its reputation or cost him a business deal. Contacting the users to get the required information, gives a very good insight in the real functioning of the product. But users have no insight in the internal production process. Users do not know if the product complies to privacy legislation or in what way the process is secured, e.g. by having more than one employee who knows the whole production process. So contacting both parties would give the best insight, but is very time consuming.
Although variations on this method can be researched (chapter 5.3), **the method of using a combination of open and closed questions to gather information about a product and to compare the results to a baseline to select the product that is 'fittest for the purpose', is indeed usable**.

Sub-question four: To what extent can the producer of NI positively influence these aspects?
Although the producer is dependent on others for part of his production process, he still has a lot of influence on the different aspects and therewith on the usability of his product.
The influence depends on the way the producer has designed the process from buying the raw data to delivering the product to the user. The more the producer can arrange internally, the more influence he has on the process and the product. But some parts simply have to be outsourced. Especially with buying the raw data, the producer is as much a client, as his clients are to him, with limited influence on the process of the data provider.
These limitations aside, adopting (inter)national standards is a good starting point. This would benefit transparency towards the user and optimize the use because of interchangeability of the NI with other information products.
Probably the most important thing for the producer to do, is actively contact both existing and potential clients. Only the users can specify their needs and complying to these needs is the basis of maximal usability.

5.3 Discussion

5.3.1 Reflecting on the research method

When starting this research I assumed that all NI products would basically be the same, at least that the aspects of usability would apply in more or less the same way to all products and therewith all products could be compared to each other, based on the same set of rules. This is not the case.
During the search for the definition of usability, I soon found that the user and his needs play a central part in assessing the degree of usability of a NI product.
I still think that the aspects I indicated to be of influence on usability, apply to all NI products, but that the requirements depend on the situation. So for every unique situation, the questions and the baseline, to which the products are compared to find the product that is most suited for that specific situation, have to be tailored to the needs of the user.

During the interviews, I experienced that for answering the questions, often two persons are needed, because in most organizations responsibilities on data and service are divided. This means that the interviewer has to be very clear about what he is expecting of the participants and that the participants have to prepare the interview.

Probably the biggest weak spot in this research is the dependency on the openness and honesty of respondents (the producers of NI). It is assumed that they have truthfully answered the survey questions, but for a number of questions there is no way of checking this. When questions are answered untruthfully, it has an effect on the outcome of the study.
During the interviews I noticed that organizations, be it governmental or commercial, are very willing to give information about their product, even if the information is not very positive. But it is always good to remember that the answers that the producers give are not (completely) verifiable, and might not be correct or complete.
Involving users in the research, would give more certainty on the validity of a part of the information. A measure that was used in this research, was the incorporation of open questions in the questionnaire. This allows the truth of the answers to be tested to a certain extent, because the participant has to elaborate on an answer to a closed question. For example, after stating that a formal security architecture is implemented, the participant is asked which architecture it is. Although this still does not guarantee the truthfulness of the answer, it is at least a step in the right direction.

5.3.2 Variations on the research method

Because I used different types of NI, I compared apples to oranges. For example, I included both governmental and commercial suppliers, products with a huge amount of users versus niche products, paid versus free and, most important, products with different purposes.
In practice, at least the purpose of the NI products will most likely be the same (e.g. flood prediction or routing) and with that, probably the type of data as well (e.g. personal information is or is not used in

creating the NI). This means that the questions and the requirements, in the form of the baseline (appendix B), should be tuned to the type of product and to the particular needs of a customer.

This would mean that for every situation, a (partially) different set of questions would be part of the questionnaire and that the baseline would be adjusted as well. This would improve the comparability of the products that are part of the research and the usability of the outcome for the user. For example, when no personal data is used in creating the product, no privacy related questions are included in the questionnaire.

Based on the interviews I conducted, I found that the interview should contain about twenty-five closed questions, including the start-up questions, with a number of open question for explanation. This would take about one hour and should provide all necessary information about the product and the producer, providing both the interviewer and the participants are well prepared.

It could be an option to create a collection of about forty to fifty questions, from which the researcher selects the twenty-five most appropriate questions. This set of questions can be used to asses a group of NI for one specific situation.
And although all questions in the interview should be relevant, not all questions have to be of equal importance. In some situations, a number of properties could be more important than other. In this situation, a weighting of the questions can be introduced.

5.3.3 Practical implications

The method that is used in this research can be used by producers to compare their product to that of the competition or it can be used by (potential) customers to find the product that best meets his needs.
It is my hope that, now that the aspects that influence the usability of NI are defined, the producers of NI will be motivated to improve the usability of their product where needed. This would cut both ways. The product would better serve the needs of the user and a satisfied customer is good for business.

5.4 Recommendations

5.4.1 Recommendations for further research

These recommendations can be divided in three groups. First are recommendations following the results of the study, second are variations on the method used and third are topics that build on and/ or expand this study.

Recommendations following the results of the study:

Based on the finding that the type of producing organization has an influence on the usability of NI products (chapter 5.2.1), what other characteristics (other than in table 1 in chapter 4) can be identified

to describe NI products and can other relations be established between product characteristics and product usability.

Another question, based on the same finding, would be whether or not government agencies pay more attention to data quality and security than companies. If yes, why? Is it data driven or does the government feel a deep responsibility towards the citizens? Could it be because of the value or sensitivity of the data, do companies really 'neglect' some parts of the process because of costs?

An extra aspect that could be important in the practical use of a NI is the ease of which the NI can be used, in other words, how easy can the user access the information. This aspect has been ignored in this study because the focus was on the usability of the information itself and not the ease with what the information can be accessed.
It is very well possible that some users find this an important issue, maybe even the most important issue, in choosing the product they are going to use. So scoring the usability of NI can be expanded with the 'ease of use' of the NI application.

Variations on the method used:
In this research, the aspects that influence the usability of NI are identified on the basis of literature review. This is a theoretical approach to usability. A qualitative research could be conducted to find out which aspects users and producers think are important for usability. Do these ideas match or do they contradict?

In collecting data, users could be the primary source of information instead of the producers. This would require new, or at least different formulated questions. In addition, both users and producers could be interviewed. The results can be compared to see if, and so, to what extent, the outcomes are different.

In a new comparison, the products that are compared should at least have the same intended use. That would provide for a realistic comparison.

Are there other ways to check the information that is given by the producers than just asking the users of the products?

Topics that build on and/ or expand this study:
This study could be implemented on a bigger scale, for different categories of NI, to gain insight in the 'state of usability' of NI on a national or even international scale. The categories could be based on type of NI, the intended use, the producer and/ or other properties of the NI.

Also, the research could be repeated over a certain amount of time. The ‘state of usability’ could evolve to a trend. Do some aspects develop quicker than other aspects (e.g. data quality), why is this? Is it because of initiatives like Inspire or because of the rappit technological developments.

With NI, communication between machines, without human interference, is becoming normal. What would be the implications on the different aspects of usability of NI. For example, to what extent are machines capable of determining whether the NI product complies to national legislation?

What is the influence of the ever increasing amount of data that is produced and that is made available? Both the government and the public create data that are, to a greater or lesser extent, made available for use by companies, individuals or the government. This is data that is being shared on social networks or that is part of the Dutch basic registrations (ICTU, 2013), that include for instance topographic information and aggregated information about crime and income.

How big is the role of social media at this moment and what are the advantages or implications of that involvement on for example privacy and to what extent are people aware.

5.4.2 Recommendations for governments

The high score of governmental agencies on quality is good for two reasons. First, the government has a (moral) obligation to provide the public with accurate information. This applies to both NI products (like Risicokaart) as to source data (like NWB and BRT). Second, data that are generated and provided by the government, are often combined with other data to create new products, be it by companies or the government itself. Errors in the source data will cause errors in the final product (error propagation).
Because of the vast amount of data the government creates and manages, and the fact that a lot of more or less autonomous agencies collect and process the data, makes initiatives like Inspire are very important. Standardization make the exchange and combination of data from different sources easier.

5.4.3 Recommendations for companies

What surprised me in the results of the interviews, was the relatively low score of companies, compared to governmental agencies, on quality and security.
From an economic point of view, it is very understandable, at least on the short term, that companies put little effort in the quality of the NI products they create and offer. Most things that have to do with quality are labour intensive and thus expensive. On the long term however, errors in the information product can cause reputational and relational damage or claims, depending on the nature and size of both product and error. Commercial organizations will always be looking for the optimal balance between quality, cost and time.

Regarding security, especially for companies of which the NI has a relatively low financial or social impact for the user, like Buienradar, it is understandable they put less effort in data security. As with

data quality, security costs time and money. But often it is the companies ICT infrastructure rather than the data a hacker is interested in. The damage then would be indirect, e.g. loss of performance.

Regarding the method used in this study, whether NI products are compared to a baseline that is setup by a (potential) user, or to a product that is “recognised as the best within the area” (Andersen and Pettersen, 1996, p.3), will depend on the purpose of the benchmark.
This research focussed on users who are interested in finding a product that best soothes their needs for a specific purpose. These users formulate their own set of conditions (a baseline) to which the NI products are compared. Based on the outcome, they then can decide which product to choose.
In practice, a producer of a NI product can very well use this method to compare his own product to other, similar products, to identify areas of improvement.

References

[01] Aalders, H., 1998. Standaardisatie van de geo-kwaliteit. In: Heres, L., ed. 1998. *Kwaliteit van geo-informatie*. Delft: Nederlandse Commissie voor Geodesie (NCG). Ch.2.

[02] Abelson, H., Ledeen, K. and Lewis, H., 2008. *Blown to bits: your life, liberty, and happiness after the digital explosion*. London: Addison-Wesley.

[03] Abelson, H., 2010. *Seductive myths about privacy*. [pdf] Retrieved March 15, 2013, from http://dig.csail.mit.edu/2010/10-mythsOfPrivacy-hal/W3C_privacy_talk_Oct_2010.pdf

[04] Andersen, B. and Pettersen, P., 1996. *Benchmarking Handbook*. London: Chapman & Hall.

[05] Anderson, S. and Küster-Filipe, J., 2003. Guaranteeing Temporal Validity with a Real-Time Logic of Knowledge. In: IEEE (Institute of Electrical and Electronics Engineers), *23rd International Conference on Distributed Computing Systems Workshops (ICDCS Workshops)*. Providence, Rhode Island, USA 19-22 May 2003. IEEE Computer Society, pp. 178-183.

[06] Australian/New Zealand Standard, 1994. *AS/NZS ISO 8402:1994 - Quality management and quality assurance - vocabulary*. 3rd ed. Wellington: Standards New Zealand.

[07] Benkler, Y., 2006. *The Wealth of Networks, how social production transforms markets and freedom*. New Haven: Yale University Press.

[08] Bradburn, N., Sudman, S. and Wansink, B., 2004. *Asking Questions*. San Francisco: Jossey-Bass.

[09] Bridgis, 2012. *Populator Walkthrough*. [online] Retrieved August 10, 2013, from http://www.youtube.com/watch?v=ENUj_UxGAbY

[10] Bridgis, 2013. [online] Retrieved August 10, 2013, from http://www.bridgis.nl/

[11] Buienradar, 2013. [online] Retrieved August 10, 2013, from http://www.buienradar.nl/

[12] Business Continuity Management Institute (BCM Institute), n.d.. *A Wiki Glossary for Business Continuity Management (BCM) and Disaster Recovery (DR)*. [online] Retrieved March 15, 2013, from http://www.bcmpedia.org/wiki/Recovery_Objectives_-_Category

[13] Chrisman, N., 1991. The error component in spatial data. In: Maguire, D., Goodchild, M., and Rhind, D. (eds.), 1991. *Geographical information systems, volume 1: Principles*. 1st ed. London: Longman, pp. 165-174.

[14] Carr, D., 2004. *Putting 40,000 Readers, One by One, on a Cover. The New York Times*. [online] Retrieved March 15, 2013, from http://www.nytimes.com/2004/04/05/business/05reason.html#h

[15] Daas, P., Ossen, S. and Arends-Tóth, J, 2009. *Framework of quality assurance for administrative data sources*. [pdf] Retrieved March 15, 2013, from http://www.statssa.gov.za/isi2009/ScientificProgramme/IPMS/0040.pdf

[16] Daas, P., Ossen, S., Vis-Visschers, R. and Arends-Tóth, J., 2009. *Checklist for the quality evaluation of administrative data sources*. The Hague: Statistics Netherlands.

[17] De By, R., ed. 2001. *Principles of geographic information systems: An introductory textbook*. 2nd ed. Enschede: The International Institute for Aerospace Survey and Earth Sciences (ITC).

[18] Delphix, 2011. *The ABCs of DaaS: Enabling Data as a Service Application Delivery, Business Intelligence, and Compliance Reporting*. [pdf] Delphix Corp: Retrieved March 15, 2013, from http://marketing.delphix.com/rs/delphix/images/DaaS_Whitepaper.pdf

[19] Dwivedi, V. and Kulkarni, N., 2008. Information as a Service in a Data Analytics Scenario. In: Institute of Electrical and Electronics Engineers (IEEE), IEEE International Conference on Web Services (ICWS). Beijing, China 23-26 September 2008. Los Alamitos: Institute of Electrical and Electronics Engineers (IEEE).

[20] Dzazali, S., Sulaiman, A. and Zolait, A., 2009. Information security landscape and maturity level Case study of Malaysian Public Service (MPS) organizations. *Government Information Quarterly*, 26(4), pp. 584-593.

[21] Ernst and Young, 2004. *Global Information Security Survey 2004*. [pdf] EYGM Limited: Retrieved March 15, 2013, from http://www.issa-motorcity.org/files/GlobalInformationSecuritySurvey2004.pdf

[22] Ernst and Young, 2012. *Global Information Security Survey: Fighting to close the gap*. [pdf] EYGM Limited: Retrieved March 15, 2013, from http://www.ey.com/Publication/vwLUAssets/Fighting_to_close_the_gap:_2012_Global_Information_Security_Survey/$FILE/2012_Global_Information_Security_Survey___Fighting_to_close_the_gap.pdf

[23] European Commission, 2013. *Inspire, Infrastructure for Spatial Information in Europe*. [online] Retrieved September fourteen, 2013, from http://inspire.jrc.ec.europa.eu/index.cfm

[24] Federal Systems Management and Integration Center (FEDSIM), 1996. *Spatial data transfer standard: Guide for Technical Managers*. [pdf] Retrieved March 16, 2013, from http://mcmcweb.er.usgs.gov/sdts/training.html

[25] Flood control 2015, 2012. *About Flood Control 2015*. [online] Retrieved August 10, 2013, from http://www.floodcontrol2015.com/about-flood-control-2015

[26] GBO provincies, 2013. [online] Retrieved August 10, 2013, from http://www.gbo-provincies.nl/index.html

[27] Google, 2013. *Google play*. [online] Retrieved August 11, 2013, from https://play.google.com/store/apps/details?id=com.supportware.Buienradar&hl=nl

[28] Gray, E., 2004. *Doing Research in the Real World*. London: SAGE Publications.

[29] Grontmij, 2013. [online] Retrieved August 10, 2013, from http://www.grontmij.nl

[30] Halls, P., 2001. Geographic information science: Innovation driven by application. *Computers, Environment and Urban Systems*, 25(1), pp. 1-4.

[31] ICTU, 2013. *Eén digitale overheid: betere service, meer gemak*. [online] Retrieved September fourteen, 2013, from http://e-overheid.nl/onderwerpen/stelselinformatiepunt/stelsel-van-basisregistraties

[32] Kadaster, 2013. [online] Retrieved August 10, 2013, from http://www.kadaster.nl/web/show

[33] Kagal, L. and Abelson, H., 2010. *Access Control is an Inadequate Framework for privacy protection*. [pdf] Retrieved March 15, 2013, from http://www.w3.org/2010/api-privacy-ws/papers/privacy-ws-23.pdf

[34] Kaufmann, M. and Hintergräber, B., 2012. *IRIS-Europe II: Implementation of River Information Services in Europe*. [pdf] Retrieved March 15, 2013, from http://www.afdj.ro/download/IRIS-Europe_II_Final_Technical_Report_part_E_Activity-4_v1p0_final.pdf

[35] Koot, M., 2012. *Measuring and Predicting Anonymity*. [pdf] Retrieved April 19, 2013, from http://staff.science.uva.nl/~delaat/pubs/2012-t-1.pdf

[36] Logan, P. and Clarkson, A., 2005. Enhancing Information Security. In: Khosrow-Pour, M. ed. Information Resources Management Association International Conference. San Diego, USA 15-18 May 2005. Hershey, PA: Idea Group Publishing.

[37] Long, S., Roberge, L., Lamicela, J. and Murugesan, M., 2004. *Balancing Data Quality Against Time and Money Constraints: Tricks from the Trenches*.[pdf] Retrieved October 13, 2013, from http://www2.sas.com/proceedings/sugi29/094-29.pdf

[38] Longley, P., Goodchild, M., Maguire, D. and Rhind, D., 2005. *Geographical Information Systems and Science*. 2nd ed. Chichester: John Wiley & Sons.

[39] McCullagh, D., 2004. Database Nation: The upside of "zero privacy". Reason, June 2004 Issue, pp. 26-35.

[40] Microsoft, 2013. *Understanding Downtime*. [online] Retrieved March 15, 2013, from http://technet.microsoft.com/en-us/library/aa998704(v=exchg.65).aspx

[41] Nieuwland, 2009. [online] Retrieved August 10, 2013, from http://www.nieuwland.nl

[42] Nieuwland, 2013. [online] Retrieved August 10, 2013, from http://geo.nieuwland.nl/index.cfm?fuseaction=news.dsp_news&recordid=194

[43] Open Geospatial Consortium (OGC), 2013. *GeoPackage Standards Working Group*. [online] Retrieved March 15, 2013, from http://www.opengeospatial.org/projects/groups/geopackageswg

[44] Ott, T. and Swiaczny, F., 2001. *Time-integrative Geographic Information Systems: Management and Analysis of Spatio-Temporal Data*. New York: Springer.

[45] Overheid.nl, 2013a. *Wet- en regelgeving, Kadasterwet.* [online] Retrieved August 31, 2013, from http://wetten.overheid.nl/BWBR0004541/Hoofdstuk1a/Titel2/Artikel7k/geldigheidsdatum_31-08-2013

[46] Overheid.nl, 2013b. *Wet- en regelgeving, Wet openbaarheid van bestuur.* [online] Retrieved October 13, 2013, from http://wetten.overheid.nl/BWBR0005252/geldigheidsdatum_13-10-2013

[47] Oxera, 2012. *What is the economic impact of Geo services* [pdf] Retrieved April 9, 2013, from http://www.oxera.com/Oxera/media/Oxera/downloads/reports/What-is-the-economic-impact-of-Geo-services.pdf

[48] Oxford University Press, 2012. *Oxford dictionaries*. [online] Retrieved January 9, 2013, from http://oxforddictionaries.com/definition/english/research%2Band%2Bdevelopment?q=research+and+development

[49] PDOK, 2013a. [online] Retrieved August 10, 2013, from https://www.pdok.nl

[50] PDOK, 2013b. *Gebruiksrapportage PDOK juni 2013*. [pdf] Retrieved August 10, 2013, from https://www.pdok.nl/nl/actueel/nieuws/artikel/18jul13-gebruiksrapportage-juni-2013-beschikbaar

[51] PDOK, n.d.. *Gebruikers aan het word*. [online] Retrieved August 10, 2013, from https://www.pdok.nl/nl/hoe/pdok-de-praktijk/gebruikers-aan-het-woord

[52] Rajesh, S., Swapna, S. and Shylender Reddy, P., 2012. Data as a Service (DaaS) in Cloud Computing, Data-As-A-Service in the age of data. *Global Journal of Computer Science and Technology Cloud & Distributed*, 12(11), pp.25-29.

[53] Rauch, H., 2012. *The Netherlands' iStrategy*. [pdf] Retrieved October 13, 2013, from http://www.government.nl/documents-and-publications/notes/2012/03/30/the-netherlands-istrategy.html

[54] Risicokaart, 2013. *Risicokaart.nl*. [online] Retrieved March 15, 2013, from http://www.risicokaart.nl/

[55] RWS, 2013. [online] Retrieved August 10, 2013, from http://www.rijkswaterstaat.nl/

[56] Saleh, M., 2011. Information Security Maturity Model. *International Journal of Computer Science and Security (IJCSS)*, 5(3), pp. 316-337.

[57] Scheuren, F., 2004. *What is a survey?* [pdf] Retrieved May 9, 2013, from https://www.whatisasurvey.info/

[58] St-Germain, R., Aliu, F., Lachapelle, E. and Dewez, P., 2012. *Societal security - Business continuity management systems*. [pdf] Retrieved March 15, 2013, from http://pecb.org/iso22301/iso22301_whitepaper.pdf

[59] Stichting Waarmerk drempelvrij.nl, 2012. *Waarmerk drempelvrij.nl*. [online] Retrieved August 10, 2013, from http://www.drempelvrij.nl/

[60] Tortorella, M. and Driscoll, P., 2005. *Reliability of information-fueled services in network-centric operations*. [pdf] Retrieved March 2, 2013, from http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA464213

[61] Trochim, W., 2006. *Research methods Knowledge Base*. [online] Retrieved May 9, 2013, from http://www.socialresearchmethods.net/kb/index.php

[62] Unigis education, 2012. *Unigis Amsterdam*. [online] Retrieved January 9, 2013, from http://www.unigis.nl/contents/unigis_amsterdam/home.asp?vb_title=home

[63] United Nations General Assembly, 1948. *The Universal Declaration of Human Rights (UDHR)*. [online] Retrieved March 2, 2013, from http://www.un.org/en/documents/udhr/index

[64] User Experience Professionals' Association (UXPA), 2012. *Usability Body of Knowledge.* [online] Retrieved March 15, 2013, from http://www.usabilitybok.org/

[65] Van Oort, P., 2005. *Spatial data quality: from description to application*. Delft: Nederlandse Commissie voor Geodesie (NCG). PhD. Koninklijke Nederlandse Akademie van Wetenschappen.

[66] Van Poperingen, J., 2004. *Kwaliteit van Geo-informatie in theorie en praktijk*. MSc. Vrije Universiteit Amsterdam.

[67] Warren, S. and Brandeis, L., 1890. *The right to Privacy*. [online] Retrieved March 16, 2013, from http://web.archive.org/web/20110429202545/http://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand_warr2.html

[68] Welchel, N., 2003. S*urvey research*. Raleigh: North Carolina State University.

[69] Williams, P., 2012. *The New Wave of Data as a Service* [online] Retrieved January 9, 2013, from http://www.dataversity.net/the-new-wave-of-data-as-a-service/

[70] Woudenberg, M., 2012. *Wettelijke grondslag verplicht gebruik basisregistraties.* [online] Retrieved August 31, 2013, from https://wiki.stelselvanbasisregistraties.nl/xwiki/bin/view/Stelselhandboek/Wettelijke+grondslag+verplicht+gebruik+basisregistraties

[71] Working Council for Chief Information Officers (WCCIO), 2003. *Trends in Information Security and Business Continuity Planning*. [pdf] Retrieved March 15, 2013, from http://m0m0g33k.dyndns.org/CEB%20Trends%20in%20Information%20Security.pdf

Appendix A

Questionnaire including cover letter and instructions

Dear [Mr/ Ms & name],

To complete my MSc GIS management course from Unigis at the Free University Amsterdam, I am conducting a research. The topic is a study to the usability of Networked Information (NI) and how this might be measured. The definition of NI is that it is generated real-time from at least one source available via the internet and it is distributed online. It is a result of a Model as a Service (MaaS), which consists of the elements data, service and application.

VI is often used to support research and decision making, e.g. with spatial planning and to identify weak spots in development plans regarding disaster management, in which combining locations with the population density can help in creating evacuation plans.
During this research, it is assumed that usability of NI, is influenced by five aspects, namely: privacy, quality of data, reliability and security of the service and the transparency of the producer.
In practise, this method will be used by potential clients who need a NI product for a specific purpose. They will want to objectively and equally compare multiple NI products that can serve the intended purpose, and choose the one that best suits their needs.

To test this theory I need information from the work field. A total of seven producers of different products are interviewed, thus [name product] is the only product that [goal of the product in short]. The research focuses on NI in general and not one specific type of product e.g. route information, weather forecast or flood prediction.

Your product is used by [discription of target audience] to [goal of the product described in more detail]. With that, [name product] has [social, financial, etc.] impact on [society, management of ..., etc.]. Also, [name producer] is [an established company in the field of … and thus has a lot of experience OR a new, innovative player in the field of …] and therefore your input is very valuable to me.

Participating in this interview could provide [name company] a (renewed) awareness about aspects that influence the usability of the product and thus provide points if improvement or confirm that [name company] is already performing very well on this topic.

The interview consists of 27 questions and takes an hour to complete. I would like to ask you to prepare the interview by going through the questions and collect information that is not readily available.

You will not be mentioned by name in the report, the product and the company, however, will. The result of this interview will be compared to results of other interviews by means of a benchmark.
With your permission, the summary of this interview (not in combination with information about other products) could be presented to one or more users of [name product] so they can comment on the outcome.

Should you need any further information, please do not hesitate to contact me by phone (06 440 589 76) or e-mail (wouter.konst@gmail.com). I look forward to seeing you on [day and date] at [time] in [address].

Kind regards,
Wouter Könst

PART 1

Introduction

To be completed by interviewer:

Name company: ……..
Name product: ……..
Location: ……..
Date: ……..
Starting time: …… End time: ……....

The interview consists of three parts. Part one (this section) provides information on terms used. In part two, questions are included that help describe the product. In part three, questions are substantive, for example, in what way the security of the product is handled.
All questions are specific for [product name].

The interview consists of 27 questions and takes an hour to complete. I would like to ask you to prepare the interview by going through the questions and collect information that is not readily available.

Technical definitions:

Data lineage:	Information on sources, update activity with dates, and processing steps that have transformed the data.
Quality:	In this thesis, quality is characterised by correctness, consisting of 'positional accuracy' and 'attribute accuracy', completeness and temporal accuracy.
Metadata:	Description of a specific dataset or service.
Personal information:	Consists of information about a natural person and that person is identifiable.
Networked information:	NI is generated real-time from at least one source available via the internet and it is distributed online.
NI:	See networked information.

General definitions:

User:	Person or organization who buys or uses the product. An organization with multiple users is seen as one user.
Participant:	Participant in the interview.
Organization:	Producer of the product.
Product:	The information product that is provided by the producer

.

PART 2

Product description

1. The participants position is:

   .....

2. The intended use of the product is:

   .....
   .....
   .....
   .....

3. The number of users of the product is:

   .....

4. The two main users types of the product are:
   - [ ] Municipalities
   - [ ] Provinces
   - [ ] State
   - [ ] Emergency Services
   - [ ] Companies
   - [ ] NGO's
   - [ ] Individuals

5. What are the three biggest customers at this time

   .....
   .....
   .....
   .....
   .....
   .....

6. Customer satisfaction is being gauged:
   - [ ] Yes, on the basis of user groups
   - [ ] Yes, on the basis of a periodical survey
   - [ ] Yes, this has been researched
   - [ ] Yes, other (please specify) .....
   - [ ] No, it is not (*skip to 8*)

7. The results may be used in this study:

   [ ] Yes

   [ ] No (please specify) .....

8. User costs are calculated on the basis of:

   [ ] The amount of time the product is used (e.g. hours per month).

   [ ] The quatity of used product (e.g. area or number of units).

   [ ] A fixed price.

   [ ] Not at all, use of product is free of charge.

   [ ] Other (please specify) .....

PART 3

Organizational questions

Privacy

Does the Personal Data Protection Act apply to the product (or parts of the product) and if so, whether it is respected. When applicable, it is examined whether the organization is complying to privacy legislation, using open follow-up questions.

9. Personal data are part of the product.
   - [ ] Yes, it is
   - [ ] No, it is not
   - [ ] Unknown

10. Personal data are part of the one or more data resources used in creating the product.
   - [ ] Yes, it is
   - [ ] No, it is not
   - [ ] Unknown

(*Questions 9 and 10 are both answered negative or unknown, skip to 12*)

11. The Personal Data Protection Act is complied.
   - [ ] Yes, fully
   - [ ] Yes, partially
   - [ ] No, it is not
   - [ ] Unknown

Following questions 9, 10 and 11:

- In what way are personal data part of the product?
- In what way are the data anonimized? What method is used?
- Who performs the operation?

Quality

The purpose of following questions is to figure out how good the product is documented, whether the organization focusses on quality and if the organization has enough confidence in the product to guarantee the quality of the product. In this study, quality consists of completeness, correctness and temporal accuracy.

12. Data quality is regularly checked
    [ ] Yes, it is
    [ ] No, not regularly
    [ ] No, data quality is not checked at all (*skip to 14*)

13. Quality checks are automated.
    [ ] Yes, fully
    [ ] Yes, partially
    [ ] No, data checks are performed manually

Following questions 12 and 13:

- To what detail is the verification procedure described?
- How often is the procedure revised?

14. Within the organization, quality management for the product focusses on:
    (*multiple answers are possible*)
    [ ] Completeness
    [ ] Correctness
    [ ] Temporal accuracy
    [ ] None, no quality management is active (*skip to 16*)

15. Within the organization, quality management for the product is secured on the level of:
    [ ] Employee
    [ ] Team leader
    [ ] Head of the department
    [ ] Director

Following questions 14 and 15:

- Which key performance indicators (KPI's) are applied?
- Can you give one or more examples?
- How often are the results reported to the person responsible?

16. The objects and attributes in the dataset are documented in a data model.
    [ ] Yes, fully
    [ ] Yes, partially
    [ ] No, they are not

Following question 16:

- How often is the data model revised?

17. The data lineage of the networked information is documented.
    [ ] Yes, fully
    [ ] Yes, partially
    [ ] No, it is not

18. The metadata (standards ISO 19115 concerning geography and ISO 19119 concerning services) of the networked information is documented.
    [ ] Yes, fully
    [ ] Yes, partially
    [ ] No, it is not

Following questions 17 and 18:
- Who documents data lineage and metadata?
- Where is the information documented?
- Who is interested in this information?

19. The organization guarantees the quality of the product.
    (*multiple answers are possible*)
    [ ] Yes, completeness
    [ ] Yes, correctness
    [ ] Yes, temporal accuracy
    [ ] No, the quality is not guaranteed

Following question 19:
- Why is the quality (not) guaranteed?

Reliability

The purpose of these questions is to determine the reliability of the service. The focus is on both the organization and the technique.

20. The production process of the product is known within the organization.
    (*multiple answers are possible*)
    [ ] Yes, the entire process is known by multiple people
    [ ] Yes, different parts of the process are known by multiple people
    [ ] Yes, the entire process is known by one person
    [ ] Yes, one or more parts of the process are known by one person
    [ ] No, it is not

Following question 20:

- YES > In what way is the knowledge about the process being shared among the employees?
- NO > In what way is the continuity guaranteed if (part of) the knowledge is not present within the organization?

21. After discontinuation of the organization (e.g. because of bankruptcy), the continuity of the product is guaranteed.
    [ ] Yes, longer than half a year
    [ ] Yes, up to half a year
    [ ] No, it is not
    [ ] Unknown

Following question 21:

- YES > In what way is the continuity secured?
- YES > How is this offered to the client?
- YES or NO > In what way is the client informed about this subject?

22. The best guaranteed server uptime in 2012 was (according to SLA or another form of contract):
    [ ] > 99,5%
    [ ] 99% to 99,5%
    [ ] 98% to 99%
    [ ] < 98%
    [ ] Unknown

23. The achieved server uptime in 2012 met the guaranteed value.
    [ ] Yes, it did
    [ ] No, it did not
    [ ] Unknown

Security

What structural measures has the organization taken to ensure security of the product and on what level is this secured?

24. The organization has implemented a formal security architecture framework (e.g. the 'open group architecture framework' or 'ANSI/IEEE 1471:ISO/IEC 42010').
    [ ] Yes, it has
    [ ] No, but it is planned for 2014
    [ ] No, it has not
    [ ] Unknown

Following question 24:

- YES > What architecture is chosen?
- YES > Why is this architecture chosen?
- NO > For what reason is this not (yet) done?
- NO > What security measures have been taken?

25. The organization has an active security management.
    (*multiple answers are possible*)
    [ ] Yes, focussed on external threads
    [ ] Yes, focussed on internal threads
    [ ] No, the organization has not (*skip to 27*)
    [ ] Unknown (*skip to 27*)

26. Within the organization, security management is secured on the level of:
    [ ] Employee
    [ ] Team leader
    [ ] Head of the department
    [ ] Director

Following questions 25 and 26:

- Which key performance indicators (KPI's) are applied?
- Can you give one or more examples?
- How often are the results reported to the person responsible?

27. Security risks have been identified.
    [ ] Yes, they have
    [ ] No, but this is planned for 2014
    [ ] No, they have not
    [ ] Unknown

Following question 27:

- YES > What do you think is the biggest risk?
- NO > For what reason is this not (yet) done?

Appendix B

Baseline for part 3 of the questionnaire and aspect ‘transparency’

The information that comes from part three of the questionnaire (Appendix A) is used to assess the usability of the different products. All answers have a predefined base score, with which the benchmark in chapter 4 (table 10) is provided.
The scores are expressed in terms of + (above baseline), o (on baseline) and - (below baseline).

9. Personal data are part of the product.

Answers	Scores	Notes
Yes, it is	nvt*	* Score depends on the answer to Q 11
No, it is not	o**	** In combination with 'No' as an answer to Q 10
Unknown	-***	*** In combination with 'Unknown' as an answer to Q 10

10. Personal data are part of the one or more data resources used in creating the product.

Answers	Scores	Notes
Yes, it is	nvt*	* Score depends on the answer to Q 11
No, it is not	o**	** In combination with 'No' as an answer to Q 9
Unknown	-***	*** In combination with 'Unknown' as an answer to Q 9

11. The Personal Data Protection Act is complied.

Answers	Scores	Notes
Yes, fully	o	The Personal Data Protection Act must be fully respected
Yes, partially	-	Partially obiding the law is still punishable
No, it is not	-	Does not comply to the law
Unknown	-	To not know is even worse

12. Data quality is regularly checked.

Answers	Scores	Notes
Yes, it is	+	Regular checks improves insight in data quality
No, not regularly	o	Checks are being performed
No, data quality is not checked at all	-	No quality checks means no insight in quality

13. Quality checks are automated.

Answers	Scores	Notes
Yes, fully	+	Improves consistency in checks
Yes, partially	o	Mix of consistency and human insight
No, data checks are performed manually	-	Increases chance of human error

14. Within the organization, quality management for the product focusses on: (*multiple answers are possible*)

Answers	Scores	Notes
Completeness	+/o/-*	* a combination of the three is desirable (+), when on element is missing, the score is (o), when two elements are missing, the score is (-)
Correctness	+/o/-*	
Temporal accuracy	+/o/-*	
None, no quality management is active	-	Unacceptable

15. Within the organization, quality management for the product is secured on the level of:

Answers	Scores	Notes
Employee	-	Insufficient
Team leader	-	Insufficient
Head of the department	o	Quality management should anchored in the organization with a certain weight
Director	+	Emphasises the importance

16. The objects and attributes in the dataset are documented in a data model.

Answers	Scores	Notes
Yes, fully	o	The size and content of the product should be fully transparent
Yes, partially	-	Insufficient
No, they are not	-	Unacceptable

17. The data lineage of the networked information is documented.

Answers	Scores	Notes
Yes, fully	o	The different parts of and the operations on the product should be fully transparent
Yes, partially	-	Insufficient
No, they are not	-	Unacceptable

18. The metadata (standards ISO 19115 concerning geography and ISO 19119 concerning services) of the networked information is documented.

Answers	Scores	Notes
Yes, fully	o	Information about the product should be fully transparent
Yes, partially	-	Insufficient
No, they are not	-	Unacceptable

19. The organization guarantees the quality of the product.
(*multiple answers are possible*)

Answers	Scores	Notes
Yes, completeness	o/-*	* a combination of the three is desirable (o), when on element is missing, the score is (-)
Yes, correctness	o/-*	
Yes, temporal accuracy	o/-*	
No, the quality is not guaranteed	-	

20. The production process of the product is known within the organization.
(multiple answers are possible)

Answers	Scores	Notes
Yes, the entire process is known by multiple people	+	If one employee would leave the organization, the production process would continue
Yes, different parts of the process are known by multiple people	o	
Yes, the entire process is known by one person	-	Insufficient
Yes, one or more parts of the process are known by one person	-	
No, it is not	-	Unacceptable

21. After discontinuation of the organization (e.g. because of bankruptcy), the continuity of the product is guaranteed.

Answers	Scores	Notes
Yes, longer than half a year	+	This would give the client sufficient time to look for alternatives
Yes, up to half a year	o	
Nee	-	Unacceptable
Onbekend	-	

22. The best guaranteed server uptime in 2012 was (according to SLA or another form of contract):

Answers	Scores	Notes
> 99,5%	+	The more the better
99% to 99,5%	o	The client is dependent on the product and so it should be easily accessible.
98% to 99%	o	
< 98%	-	Insufficient
Unknown	-	Unacceptable

23. The achieved server uptime in 2012 met the guaranteed value.

Answers	Scores	Notes
Yes, it did	o	As it should be
No, it did not	-	Insufficient
Unknown	-	Unacceptable

24. The organization has implemented a formal security architecture framework (e.g. the 'open group architecture framework' or 'ANSI/IEEE 1471:ISO/IEC 42010').

Answers	Scores	Notes
Yes, it has	+	
No, but it is planned for 2014	o	It shows the organization is conscious about security
No, it has not	-	Insufficient
Unknown	-	Unacceptable

25. The organization has an active security management.
(multiple answers are possible)

Answers	Scores	Notes
Yes, focussed on external threads	o/-*	* a combination of the three is desirable (o), when on element is missing, the score is (-)
Yes, focussed on internal threads	o/-*	
No, the organization has not	-	Insufficient
Unknown	-	Unacceptable

26. Within the organization, security management is secured on the level of:

Answers	Scores	Notes
Employee	-	Insufficient
Team leader	-	
Head of the department	o	Security management should anchored in the organization with a certain weight
Director	+	Emphasises the importance

27. Security risks have been identified.

Answers	Scores	Notes
Yes, they have	o	
No, but this is planned for 2014	o	It shows the organization is conscious about security
No, they have not	-	Insufficient
Unknown	-	Unacceptable

Transparency		
Topic	**Scores**	**Notes**
Interview	o	If the participant of the interview is cooperating willingly
	-	If data was purposely withheld
Contact	+	If one or more contacts are named (personaly)
	o	If a contact address is given (at least mail)
	-	No contact information
Used sources	+	Information about all sources is given
	o	Information about the main sources is given
	-	No source information is given
Intended use	o	Intended use is given
	-	Intended use is not given
User costs	o	User cost use is given
	-	User cost use is not given
Manual	+	Extensive manual
	o	Short manual
	-	No manual
References	o	References are given
	-	References are not given

Appendix C

Answers given in part 3 of the questionnaire

The answers of part three of all questionnaires (Appendix A) have been combined. For each question it is indicated how many times an answer is given.

Other than seven answers per question can occur when a participant has to skip a question or is allowed to give multiple answers per question. The premise is that 7 responses is 100%.

9. Personal data are part of the product.

Answers		Response total	Response percentage
Yes, it is			
No, it is not		7	100%
Unknown			
	Total respondents:	7	100%

10. Personal data are part of the one or more data resources used in creating the product.

Answers		Response total	Response percentage
Yes, it is		1	14%
No, it is not		6	86%
Unknown			
	Total respondents:		100%

11. The Personal Data Protection Act is complied.

Answers		Response total	Response percentage
Yes, fully		1	14%
Yes, partially			
No, it is not			
Unknown			
	Total respondents:	1	14%

12. Data quality is regularly checked.

Answers		Response total	Response percentage
Yes, it is		3	43%
No, not regularly		3	43%
No, data quality is not checked at all		1	14%
	Total respondents:	7	100%

13. Quality checks are automated.

Answers	Response total	Response percentage
Yes, fully		
Yes, partially	3	43%
No, data checks are performed manually	3	43%
Total respondents:	**6**	**86%**

14. Within the organization, quality management for the product focusses on: (*multiple answers are possible*)

Answers	Response total	Response percentage
Completeness	3	43%
Correctness	4	57%
Temporal accuracy	5	71%
None, no quality management is active	2	29%
Total respondents:	**14**	**200%**

15. Within the organization, quality management for the product is secured on the level of:

Answers	Response total	Response percentage
Employee	1	14%
Team leader	2	29%
Head of the department		
Director	2	29%
Total respondents:	**5**	**72%**

16. The objects and attributes in the dataset are documented in a data model.

Answers	Response total	Response percentage
Yes, fully	6	86%
Yes, partially	1	14%
No, they are not		
Total respondents:	**7**	**100%**

17. The data lineage of the networked information is documented.

Answers	Response total	Response percentage
Yes, fully	3	43%
Yes, partially	2	29%
No, they are not	2	29%
Total respondents:	**7**	**100%**

18. The metadata (standards ISO 19115 concerning geography and ISO 19119 concerning services) of the networked information is documented.

Answers	Response total	Response percentage
Yes, fully	3	43%
Yes, partially	2	29%
No, they are not	2	29%
Total respondents:	7	100%

1. De The organization guarantees the quality of the product.
(*multiple answers are possible*)

Answers	Response total	Response percentage
Yes, completeness	4	57%
Yes, correctness	5	71%
Yes, temporal accuracy	5	71%
No, the quality is not guaranteed	1	14%
Total respondents:	15	213%

20. The production process of the product is known within the organization.
(multiple answers are possible)

Answers	Response total	Response percentage
Yes, the entire process is known by multiple people	4	57%
Yes, different parts of the process are known by multiple people	2	29%
Yes, the entire process is known by one person	1	14%
Yes, one or more parts of the process are known by one person		
No, it is not		
Total respondents:	7	100%

21. After discontinuation of the organization (e.g. because of bankruptcy), the continuity of the product is guaranteed.

Answers	Response total	Response percentage
Yes, longer than half a year		
Yes, up to half a year	1	14%
Nee	6	86%
Onbekend		
Total respondents:	7	100%

22. The best guaranteed server uptime in 2012 was (according to SLA or another form of contract):

Answers	Response total	Response percentage
> 99,5%	1	14%
99% to 99,5%	2	43%
98% to 99%		
< 98%	1	14%
Unknown	3	29%
Total respondents:	7	100%

23. The achieved server uptime in 2012 met the guaranteed value.

Answers	Response total	Response percentage
Yes, it did	3	43%
No, it did not	1	14%
Unknown	3	43%
Total respondents:	7	100%

24. The organization has implemented a formal security architecture framework (e.g. the 'open group architecture framework' or 'ANSI/IEEE 1471:ISO/IEC 42010').

Answers	Response total	Response percentage
Yes, it has	2	29%
No, but it is planned for 2014	1	14%
No, it has not	4	57%
Unknown		
Total respondents:	7	100%

25. The organization has an active security management.
(multiple answers are possible)

Answers	Response total	Response percentage
Yes, focussed on external threads	4	57%
Yes, focussed on internal threads	3	43%
No, the organization has not	3	43%
Unknown		
Total respondents:	10	143%

26. Within the organization, security management is secured on the level of:			
Answers		**Response total**	**Response percentage**
Employee			
Team leader			
Head of the department			
Director		**4**	**57%**
	Total respondents:	**4**	**57%**

27. De veiligheidsrisico’s zijn in beeld gebracht.			
Answers		**Response total**	**Response percentage**
Yes, they have		**5**	
No, but this is planned for 2014			
No, they have not		**2**	
Unknown			
	Total respondents:	**7**	**100%**