Analysis of Surveys from an Enterprise Content Management System

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Information & Knowledge Management

eingereicht von

Markus Paulhart BSC.
Matrikelnummer 0325357

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung
Betreuer/in: Ass.Prof. Dr. Monika Di Angelo

Wien, TT.MM.JJJJ

_____   _____
(Unterschrift Verfasser/in)   (Unterschrift Betreuer/in)

Technische Universität Wien

A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.ac.at

Analysis of Surveys from an Enterprise Content Management System

MASTER THESIS

for the acquisition of the academic degree

Master of Science

within the scope of the studies

Information & Knowledge Management

submitted by

Markus Paulhart BSC.
matriculation number 0325357

on the
Faculty of Informatics at the Vienna University of Technology

Supervision
Supervisor: Ass.Prof. Dr. Monika Di Angelo

Vienna, TT.MM.JJJJ _____   _____
                        (Signature Author)   (Signature Supervisor)

Erklärung zur Verfassung der Arbeit

Markus Paulhart BSC., Palffygasse 3/8 1170 Wien

„Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit einschließlich Tabellen, Karten und Abbildungen, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe."

_____

Ort, Datum, Unterschrift

Danksagung

Diese Arbeit widme ich meinen Eltern, die mich, wo sie nur konnten, unterstützt und immer an mich geglaubt haben. Ohne sie wären mein Studium und diese Arbeit nicht möglich gewesen.

Ein besonderer persönlicher Dank gilt meiner besseren Hälfte Anna. Die mir in den die vergangenen Jahre durch ihre fortwährende Hilfe und Beistand es ermöglicht hat diese Arbeit ab zu schließen.

Bedanken möchte ich mich auch bei meiner Betreuerin Ass.Prof. Dr. Monika Di Angelo für die tatkräftige Unterstützung bei der Erstellung dieser Arbeit und auch der Firma Raiffeisen Bank Informatik für die Möglichkeit diese Arbeit erstellen zu können.

Abstract

Enterprise Content Management Systems like SharePoint are widely used in middle and large scale Companies. (Miles 2011) Their purpose is to provide a centralized portal for collecting, delivering and archiving data like documents, e-mails, files, knowledge etc. for the users. Such tools also support the user in doing empirical research, which the analyst can use for fully computer-based qualitative questioning method, also called surveys.

Now the purpose of the master thesis is to analyze how the collected survey data is presented and visualized to the user. Furthermore the possibility to analyze the data using an external tool is evaluated and implemented. The survey feature in SharePoint is hereby used as an example for a qualitative questioning method.

In order to select a suitable external analyzing application the four Business Intelligence tools (1) Spotfire from TIBCO, (2) QlikView from QlikTech, (3) Tableau Desktop from Tableau Software and (4) SQL Server Analysis Services (SSAS) Cubes from Microsoft have been evaluated. Eventually the implementation was performed with QlikView.

Comparing the external applications with the out of the box SharePoint survey data representation the results show that using an external tool more information can be gathered a lot easier and also presented in a more user-friendly way. On the downside the process of collecting all information needed is extra time consuming, which is not necessary with the standard built-in analyzing methods of SharePoint.

Kurzfassung

Enterprise Content Management Systeme wie SharePoint sind stark in Verwendung in Mittel- und Großunternehmen. (Miles 2011) Ihr Zweck ist es dem Benutzer ein zentralisiertes Portal zur Verfügung zu stellen, das ihre Daten sammelt, anbietet und archiviert (Dokumente, E-mails, Dateien, Wisse, etc.). Solch ein Portal kann den Benutzer bei der Erstellung von empirischen Studien helfen und bietet dem Analysten die Möglichkeit einer voll elektronischen qualitativen Fragemethode, dem Fragebogen.

Der Zweck dieser Diplomarbeit besteht darin, die mit einem Fragebogen erhobenen Daten nach der Art ihrer Abrufung und Darstellungsmöglichkeiten zu analysieren. Des Weiteren wir die Möglichkeit, diese Daten mithilfe eines externen Programms zu analysieren evaluiert und realisiert. Hierfür werden Fragenbögen welche mit SharePoint erstellt worden sind als exemplarisches Beispiel herangezogen.

Um ein passendes externes Analyseprogramm zu wählen wurden die folgenden vier Business Intelligence Tools, nämlich (1) Spotfire von TIBCO, (2) QlikView von QlikTech, (3) Tableau Desktop von Tableau Software und (4) SQL Server Analysis Services (SSAS) Würfel von Microsoft evaluiert. Schlussendlich wurde das Programm mit QlikView realisiert.

Wenn man die externe Applikation mit den standardisierten SharePoint Funktionalitäten zur Darstellung von Fragebögen vergleicht stellt man fest, dass man mit der externen Applikation mehr Informationen bei gleichzeitig ansprechenderer Darstellung erreichen kann. Der Nachteil ist jedoch der zusätzliche zeitliche Aufwand durch die Datensammlung, der bei der Standard SharePoint Analysemethode nicht notwendig ist.

Table of Contents

1 Introduction

1.1 Problem statement

Starting with the first official discussion in October 2004, about WEB 2.0, there was a strong realization that internet users are increasing their contribution to creating new data. Before this time, called WEB 1.0, users were mainly consuming data which was provided by specific sources, without contributing anything themselves. By changing this aspect the user switched from his role as a passive consumer to an interactive and collaborative contributor to the internet. (O'Reilly 2009)

This effect was also noticed in the business environment. Companies with classical Enterprise Content Management Systems (ECM or Enterprise CMS) recognized the change and integrated more of the so called WEB 2.0 features, for example wikis. This progress increased the number of "emerged collaboration" companies (Morgen 2012), meaning that such companies are more collaborative in their employee-interactions.

Exactly this interaction with focus on analyzing the employee's feedback, gathered by special surveys, is the starting point in this master thesis. Surveys provide great means for acquiring opinions and feedback to certain topics. SharePoint is one of the best known ECM tools on the market and is therefore picked as a reference tool in this thesis. SharePoint supports the feature of surveys quite well, but provides insufficient tools for data analysis. Data collection is only half of the way, analyzing the gathered data is just as important.

How can SharePoint survey data be analyzed in a fast, comprehensive and most importantly facile way, especially by people working in a business environment?

Referring to people in the business environment means the need to provide analysis tools for users without any specialized knowledge or additional training. Subsequently the following question emerges as well.

Under which circumstances should the survey analyst use an external analyzing tool?

1.2 Expected result

The aim of the master thesis is to look at the possible analysis of ECM survey data. In order to accomplish a feasible statement, Microsoft SharePoint 2010 has been chosen as the Enterprise CMS tool for the test scenario. Because of the following two major reasons, SharePoint was considered the best choice for representing the ECM tools on the market:

1. It is the most widely used ECM. A survey created by AIIM (Association for Information and Image Management (AIIM, About)) in February 2011 showed that out of 650 companies 58 % of them have Microsoft SharePoint installed. This number increases to 70 %, considering only companies with more than 5000 employees. (Miles 2011)

2. Two famous market research companies, Forrester Wave and Gartner, consider Microsoft with the tool SharePoint as one of the most promising companies regarding Enterprise Content Management. (Gilbert, et al. 2012) (Weintraub, Le Clair and McKinnon 2013)

Previous experiences have shown that SharePoint does not offer many elaborated ways to analyze survey data. Therefore the data will be exported out of SharePoint and analyzed using an external analysis software .The goal is to demonstrate that there are more suitable ways to analyze the data than SharePoint provides.

1.3 Structure of the master thesis

The master thesis is divided into eight chapters. Chapter 2 details the state-of-the-art in Knowledge, Content and Enterprise Management as well as Business Intelligence analyzing tools for surveys. Chapter 3 presents the methodological approach of this thesis including the implementation concept.

The practical part is structured into the following chapters:

- *Analyze SharePoint Surveys* – How SharePoint is presenting the survey data to the analyst and how data can be categorized.
- *Evaluate an external analysis tool* – Evaluation of an external analysis tool to visualize the survey data.
- *QlikView SPSurvey Dashboard solution* – Description of the QlikView application and implementation.

The concluding part consists of a summary on visualization methods of survey data using either SharePoint or an external tool followed by a conclusion shown in chapter summary & results. The last chapter provides an outlook on the analysis of survey data using external tools in the near future.

2 State-of-the-Art

This chapter sums up the scientific state-of-art regarding the following subjects of this master thesis:

- Knowledge Management
- Enterprise Content Management
- Surveys
- Business Intelligence Tools

2.1 Knowledge Management

The term Knowledge Management (KM) can be defined in many different ways. One reasonable definition is that KM is an "… interdisciplinary business model that has knowledge within the framework of an organization as its focus. It is rooted in many disciplines, including business, economics, psychology, and information management. … Knowledge management involves people, technology, and processes in overlapping parts". (Awad and Ghaziri 2007) This idea is also visualized in Figure 1.
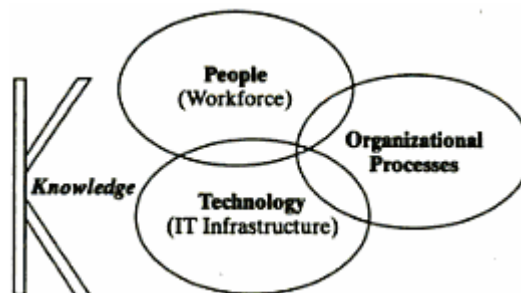
Figure 1: Overlapping Human, Organizational, and Technological Factors of KM (Awad and Ghaziri 2007)

Managing the knowledge of an enterprise or organization has always been very important in order to succeed in business. In fact, a good knowledge management can lead to an increase in performance within the company. (Griffith, Sawyer and

Neale 2003) Furthermore almost half of the enterprises stated that they installed a knowledge management system in their company to improve handling their own knowledge and not to increase the margin directly. (Chan and Chao 2008)

2.2 Enterprise Content Management

Planning to do KM in a company is one thing, knowing how to do it is a complete different one. In order to successfully do KM in a firm, ECMs are offering various ways to help with that part. ECM is a technology that collects, manages, prints, saves and archives information to support business processes within the company. Therefore the management of the content, collaboration, Web Content Management as well as the Life-Cycle management plays an important role in these products. (Eggert 2007) ECM of course also provides information for the user. For example data is collected and stored via surveys, which then can be presented or provided through the ECM.

Figure 2: ECM (OPTIMAL SYSTEMS GmbH, Definition ECM)

Figure 2 depicts a general scheme of an ECM, illustrating the data flow. Data is collected in the part called "*Erfassung*", provided in "*Bereitstellung*" and archived represented as "*Archivierung*".

Another definition by AIIM says that "Enterprise Content Management (ECM) is the strategies, methods and tools used to capture, manage, store, preserve, and deliver content and documents related to organizational processes. ECM tools and strategies allow the management of an organization's unstructured information, wherever that information exists." (AIIM 2013)

Nowadays there are many market vendors offering ECM tools. An excerpt among others of "…enterprise content and collaboration portals are Microsoft SharePoint, IBM WebSphere, Oracle WebCenter, and SAP NetWeaver." (Fenz 2012) The use of such an ECM affords the possibility, among other things, to create surveys and collect data with it.

2.3 Surveys

Horst Otto Mayer stated: "Empirische Forschung ist das Handeln nach bestimmten Regeln". (Mayer 2013), which means empirical research is acting following specified rules. In other words it is a research based on predefined plans or models to gather general valid statements on a subject from a group of people. There are two different methods for an empirical review. One is the qualitative questioning method such as a guided interview and the other is a quantitative questioning method like surveys. (Mayer 2013).

The idea of surveys is basically to ask the user predefined questions, which are answered in a digital or analog form by the user. This can be done completely computer-based, whereas interviews are hardly done without any physical human interaction and are therefore not be supported by ECM. Consequently most of the Enterprise CMS like SharePoint support only surveys as an empiric questioning method. Therefore surveys are a fast solution to gather opinions form a wide group of

people without much additional effort, if an Enterprise Content Management is already in place.

2.4 Business Intelligence tools

So far no similar attempt has been made to analyze SharePoint survey data with an external tool. The external application needs to be fast and also easy to use in order to be appropriate for an untrained business user to handle.

To fulfill both criteria a Business Intelligence (BI) tool is a suitable choice of analyzing the data in real time by using in memory computing. "In-memory computing is the storage of information in the main random access memory (RAM) of dedicated servers rather than in complicated relational databases operating on comparatively slow disk drives. In-memory computing helps business customers, including retailers, banks and utilities to quickly detect patterns, analyze massive data volumes on the fly, and perform their operations quickly." (Janssen 2013)

Figure 3 depicts the connection and interaction of a BI application (blue rectangle) with the data sources (red area) and the users.
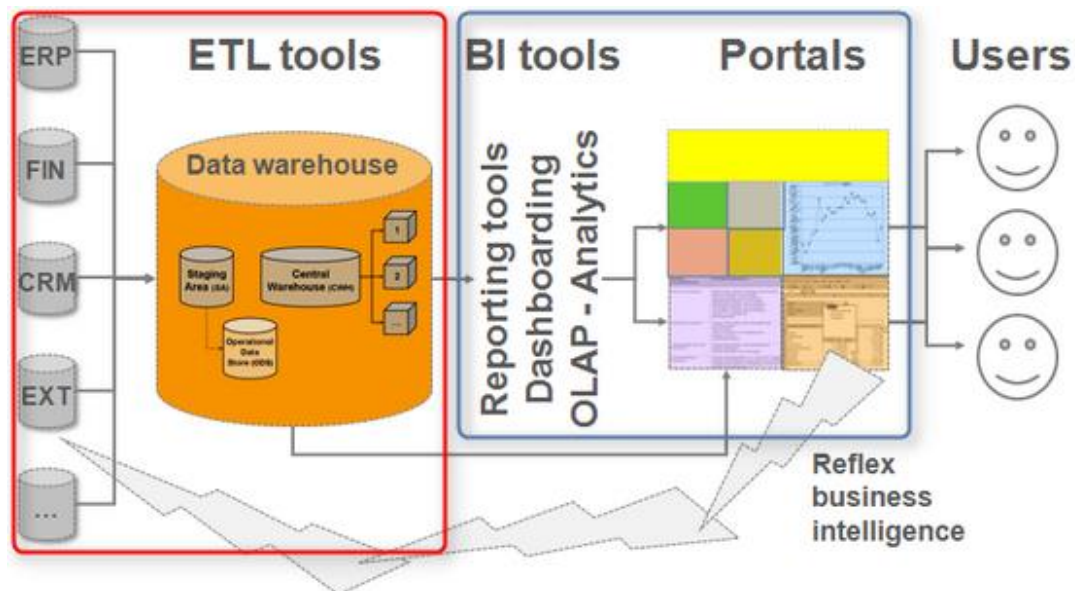
Figure 3: Interaction of a BI application (Passionned Group, Definition of BI)

Additionally Figure 3 shows that BI tools have the business user as their end customer in mind and are therefore anxious to make the software as user-friendly as possible. In comparison "…common tools such as Excel or Access … can lead to dashboards that are not very user-friendly", whereas BI tools "…take the dashboard idea to the next level. More than simply graphically" they "… can dramatically expand your insight into program performance". (Yurgosky 2012)

3 Methodological approach

This chapter defines the methodological approach for this master thesis step by step based on the following questions:

- How is the data stored in SharePoint?
- What kind of data was available for this approach?
- What are the criteria for analyzing the data?
- What kind of visualization can be generated and implemented with an external analyzing tool?

Literature regarding WEB 2.0, CMS, EMS Systems, SharePoint served as basis for the design and implementation concept of the external analysis tool.

3.1 Evaluation of SharePoint survey feature

In order to analyze the survey feature in SharePoint in more detail, we need to determine how surveys are created and which data structures are generated in the Content Management System. For this analysis no specific data set is required. On the contrary, it is sufficient to create exemplary questions for each question type in SharePoint in order to understand how data is represented and evaluated on this platform.

3.2 Evaluation of external applications for the survey data

Before designing and implementing an application that visualizes the survey data, the appropriate analysis method has to be evaluated. The analyzing tools need to work with the test data and also correspond to the analyzing criteria, which are both described below.

3.2.1 Test data

The test data for this evaluation consists of six connected surveys, done with the YPP (Young Professional Potentials) group. The YPP group is a pool of people who works in different areas in the company in order to represent prior knowledge in these designated areas. The surveys itself are about the My Site feature in SharePoint and Social Networking in a business environment. The reason for picking this dataset is because it represents a widely used set of different questions in SharePoint, which should provide a good basis for any evaluation. Therefore the external application has to be able to also visualize the exported dataset in order to be considered appropriate.

3.2.2 Criteria for the analysis

In order to analyze the external application properly three important criteria are to be examined (see below):

- Usability
- Flexible and fast analysis
- Loading the application with data without any modification of the script, programming or visualization components

3.3 Design and implementation of a visualization application

After choosing a suitable candidate the practical part of the master thesis is about the use of previously gained knowledge to create a concept for an enhancement of the analysis using an external tool. The following features need to be implemented in the application to improve the analyzing experience:

- A chart that represents all answers divided by the department including the relation of the total value.
- A detailed view on every answers given by the participant
- Filter options regarding answers, questions and question classes
- An analysis of people participating in the question or survey, also divided by their job responsibilities.

- A flexible data load, which means no additional programming or changes on the dashboard are required to load new surveys into the application.

After the implementation the application will be checked using the same criteria mentioned above.

3.4 Comparison

Finally a comparative evaluation of these two different analyzing approaches is carried out.

4 Analysis of SharePoint Surveys

The SharePoint survey is a useful and powerful tool to collect opinions and trends within a company. After data acquisition, it is even more relevant to analyze the results in a fast and flexible way. This chapter discusses what a survey can consist of and how it is represented in SharePoint.

4.1 SharePoint question types

A survey in SharePoint 2010 consists of several questions; each question can be one of 13 different types: (Microsoft, Create a survey)

- Single line of text
- Multiple lines of text
- Choice (menu to choose from)
- Rating Scale (a matrix of choices or a Likert scale)
- Number (1, 1.0, 100)
- Currency ($, ¥, €)
- Date and Time
- Lookup (information already on this site)
- Yes/No (check box)
- Person or Group
- Page Separator (inserts a page break into your survey)
- External Data (Microsoft, External Data in SharePoint 2010)
- Managed Metadata (Microsoft, Managed Metadata)

The question type *Page Separator* is not a real question. It is only used to create another page for the upcoming questions in order to maintain the readability of the SharePoint survey. It will not be considered in the further studies.

4.2 Representation of SharePoint survey data

Almost every question type listed before represents a different kind of data set. This distinctness makes the data more complex and therefore more challenging to process and analyze.

SharePoint 2010 itself offers three possibilities to analyze data collected within a survey:

- All Responses
- Graphical Summery
- Export to Spreadsheet

4.2.1 All Responses

Figure 4 depicts the first representation of all answers of one participant, responding on a specific survey.

Figure 4: mySite - Part 2 - My Profile (View Response #2)

This possibility provides a detailed perspective of all data without any calculations or aggregations. It is also designed to check single statements of participants.

4.2.2 Graphical Summery

The second data representation illustrates the results of a question grouped by answers, see Figure 5.

1. 1. How much experience do you have with SharePoint?

I'm feeling very confident with this tool and I know all features and tricks which are important for me.

3 (16%)

I think I have a good grasp of Sharepoint, but there is still room for improvement.

3 (16%)

It's ok. I just know some basic features but that's it.

10 (53%)

I am getting along with it, but I am not really able to edit or create something.

2 (11%)

I am only using it when it's necessary.

1 (5%)

Total: 19
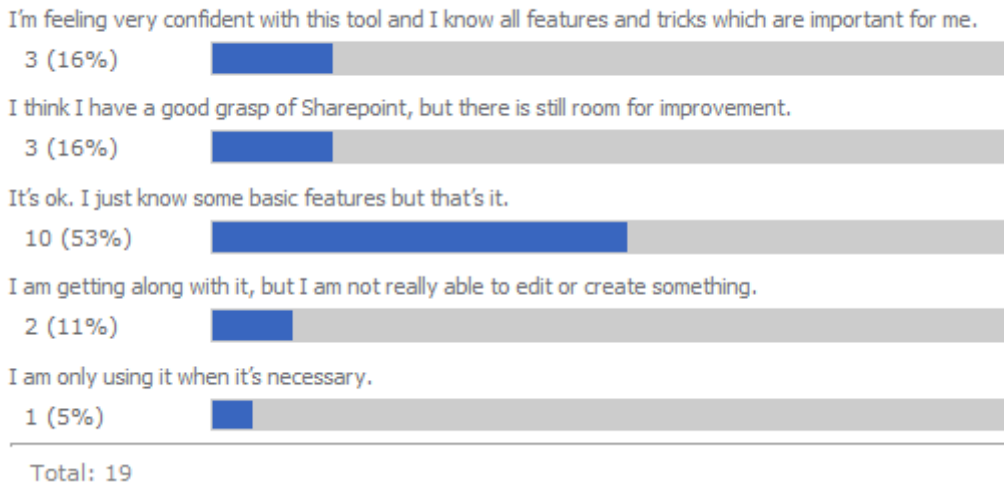
Figure 5: mySite – Getting Started – Question 1 (Graphical Summery)

Additionally three calculated numbers are included, graphically highlighted in Figure 6:

- Answers: This represents the number of people with the same answer.
- Percentage: The number of participants picking this answer in relation to the total answers value.
- Total Answers: Total number of answers on this question.

1. 1. How much experience do you have with SharePoint?

I'm feeling very confident with this tool and I know all features and tricks which are important for me.

3 (16%)

I think I have a good grasp of Sharepoint, but there is still room for improvement.

3 (16%)

It's ok. I just know some basic features but that's it.

10 (53%)

I am getting along with it, but I am not really able to edit or create something.

2 (11%)

I am only using it when it's necessary.

1 (5%)

Total: 19
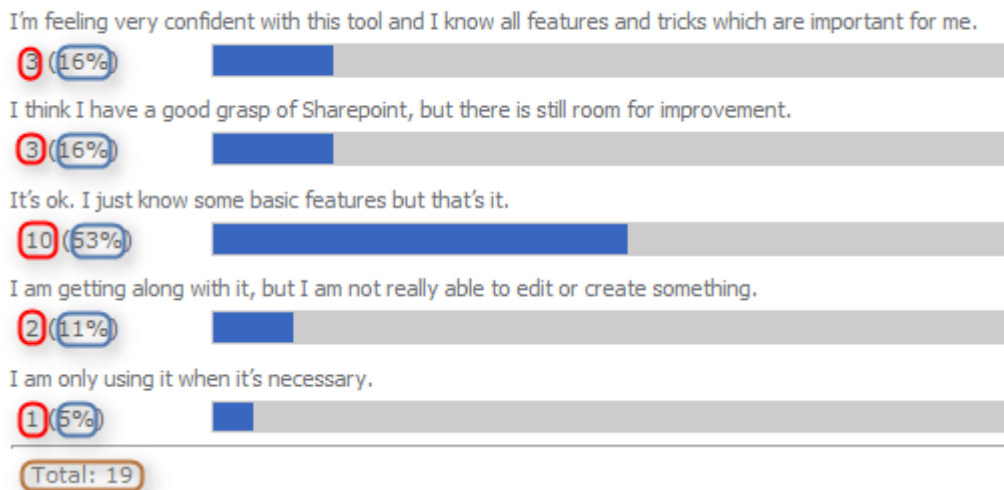
Figure 6: mySite – Getting Started – Question 1 (Graphical Summery) calculated numbers

Figure 6 shows at a look the most answered question as well as the most infrequent ones. Furthermore the analyst is able to present all given answers of all participant for one particular question including the number of the answers.

4.2.3 Export to Spreadsheet

The third representation possibility is not really a view, but enables the user to export the collected data into a spreadsheet. The spreadsheet can be opened and analyzed for example in Microsoft Excel or with an external analysis program such as QlikView by QlikTech.

This function only exports the information entered by the user, which means no further information except for the person answering and the question will be provided by SharePoint. No additional Meta data, as the type of question or the order of the answer, can be exported using this functionality. Besides, the exported data is not anonymous which needs to be considered in further analyses.

The *Export to Spreadsheet* option is the basis for all other further analysis done with an external tool.

4.3 Spreadsheet data

When it comes to analyzing the data with another application rather than SharePoint the easiest way is to import a spreadsheet version of the survey from SharePoint. As mentioned above SharePoint provides only data entered by participants, the name of the question and the answers. Therefore it is necessary to evaluate how SharePoint is presenting the data via spreadsheet in order to handle the data correctly in the analyzing tool.

Looking at the SharePoint question types it seems natural to create a unique way of representation for each type. At a second glance, the data behind the question type

seems to be similar, in some cases. Hence it makes sense to take a close look into the data itself.

In order to define the differences in the questions the goal is to analyze the answers. It turned out that some questions hold similarities with others and can be handled likewise in the analyzing application.

This understanding leads to a new summarized classification of the questions, more suitable for external analysis.

- *Plain Text*: Any kind of text can be entered.
- *Single Choice*: Only one answer is given by a participant.
- *Multiple Choices*: One or more answers are chosen by the participant.
- *Rating*: The participant can select a value related to a predefined scale.
- *Numbers*: A number, date or time can be entered.
- *Single Lookup*: Select a value belonging to an object in SharePoint.
- *Multiple Lookup*: One of more values related to an object in SharePoint can be selected.

These categories are also able to handle empty answers, which mean data without any answer selected by the participant.

The data of each category needs to be treated differently during the load, because of two reasons. (1) Each data is presented in a different structure in the spreadsheet provided by SharePoint. (2) The data needs a different presentation in the frontend and therefore requires special processing.

To facilitate the characterization, these new categories are referred to as spreadsheet question classes (SSQC) during this work.

4.3.1 Spreadsheet Question Classes

Each of the defined classes represents a different data set, which is unique regarding two aspects:

- How the data is loaded in the external analysis application.
- How the data is presented in the frontend.

According to these two aspects the analysis application needs a special design for each class in order to represent the data not only in a correct and thorough way but also to use the individual analytical capacities.

For example compared to text answers numbers allow more specific statistical analyses like calculating the average, highest lowest, sum, etc.

The following subchapters try to answer some questions regarding the SSQC:

- How does SharePoint present the data via the spreadsheet export?
- What makes the classes unique?
- What kind of SharePoint question type is covered with this class?
- How can the contained data be analyzed?

4.3.1.1 Plain Text

This class covers answers that are not easily classified. There is no predefined set of answers or other predictable logic like numbers, the answer cell could even be empty. The two major SharePoint question types covered by this class are *Single line of text* and *Multiple lines of text*.

Data stored in the Spreadsheet:
The text is presented by value fields in the corresponding spreadsheet if the participant has answered the question. Otherwise the cell is left empty. The column title contains the title of the question.

Data presented in the Frontend:

Analyzing the Plain Text class is performed by looking directly into the data. The only purpose is data representation, because the data itself is self-written text and almost always unique. Therefore no information beyond the answers is available in advance.

4.3.1.2 Single Choice

By answering a Single Choice question the participant can choose one answer of a predefined set of answers, which is not delivered via spreadsheet. It is possible for the user to provide no answer at all. The two major SharePoint question types covered by this class are *Choice (menu to choose from)* with an option to enforce unique values and *Yes/No (check box)*. Furthermore *Person or Groups*, *Managed Metadata* and *External Data* can also be allotted to this class.

Data stored in the Spreadsheet:
The text is presented the same way as in the Plain Text class. The data or the structure itself gives no hint at whether or not this data set is a Single Choice class or not.

Data presented in the frontend:
Through the fixed set of answers for each question there are more options to analyze the data. Hence it is possible to make a statement about the frequency of the answers in numbers or percentage. Whiteout a defined set of answers the visualization could add no additional value to the analysis.

4.3.1.3 Multiple Choice

The Multiple Choice class is the same as the Single Choice class with the additional possibility of more than one answer per user. The major SharePoint question type covered by this class is *Choice (menu to choose from)* with the option to allow multiple answers. Furthermore *Person or Groups*, *Managed Metadata* and *External Data* can be handled as well with this class.

Data stored in the Spreadsheet:

The data is presented as in the Single Choice except that the answers for multiple choice questions are stored differently in the Spreadsheet. Multiple answers are stored in one cell separated by *;#* like *Add Documents;#Write blogs*. This difference needs to be considered during the load of the data but also in the representation.

Regarding the detection of the question type it is possible to recognize if multivalve selection is allowed, provided that at least one of the participants has given more than one answer. Otherwise the data could also be a Single Choice SSQC.

Data presented in the frontend:
Compared to the Single Choice class the visualization needs to manage more than one answer for a participant. Because of the affiliation (for example the department of the participant) the data gains more value regarding the evaluation of the question.

4.3.1.4 Rating

Rating questions allow the user to rank the topic by choosing from a fixed set of answers, which are represented by a range of numbers.

The range can be defined in three steps:

- *poorly* - as lowest
- *more or less ok* - representing the middle value
- *good* - for the highest value

The scale and the range texts are defined by the survey creator and are not exported via the spreadsheet. Furthermore *sub-questions* can be posed. For example, the Question 3: *How strong are these functions covered in My Site. Especially in a business environment compared to Xing, Facebook, LinkedIn, etc.?* would lead to the possible sub-questions *Show your Skills*, *Keep up with friends (relationship)* or *Stay informed about your interests.* Each sub question has its own rating, but the rating settings of the main question also apply to all sub questions. Hence the major SharePoint question type covered by this class is *Rating Scale (a matrix of choices or a Likert scale).*

Data stored in the Spreadsheet:

Each sub-question is treated like a single question in the data. That means every sub-question has its own column, where the name of the column is a combination of the question and sub-question name. The cells contain only one value, which is the same as in the Single Choice class. To cover the whole question all sub-questions need to be loaded into the application. Besides, the participants can also rate a sub-question with a N/A value.

Data presented in the frontend:

The data of a sub-question can be presented the same as the Single Choice or Multiple Choice. Therefore it makes sense to analyze each sub-question separately. Furthermore it is possible to make an overall statement about how good or bad the question was rated in general.

4.3.1.5 Numbers

The number class combines a variety of figures. That is why the three major SharePoint question types covered by this class are *Number (1, 1.0, 100)*, *Currency ($, ¥, €)* and *Date and Time*.

Data stored in the Spreadsheet:

The text is presented the same as in the Plain Text class. The cells only hold numbers, dates or date/time values. They can also be empty. Most external analysis application like QlikView can handle Date or Date/Time values as numbers. This is the reason why there is no difference between numbers, date or date/time values.

Data presented in the frontend:

Depending on the questions it may be helpful to generate sum figures like average, highest, lowest, and so on. Plus presenting the values in a Scatter Chart which "… is a set of points plotted on a horizontal and vertical axes" (Rouse and Gibilisco 2012) can increase the value of the analysis.

4.3.1.6 Single/Multiple Lookup

The lookup class is basically a Single or Multiple Choice class. The answers are collected using an existing SharePoint object. Each SharePoint object has different attributes and the following will also be stored using the excel export:

- Title
- Expires
- ID
- Modified
- Created
- Version
- Title (linked to item)

Hence the major SharePoint question type covered by both classes is *Lookup (information already on this site).*

Data stored in the Spreadsheet:
The data is essentially stored the same way as in Single or Multiple Choice class. The major difference is that each additional attribute stored within the answer generates a second column in the spreadsheet. The name of the attribute columns is a combination of the question and attribute name. This makes data load in general more complex.

Data presented in the frontend:
All values, including the question and all attributes, should be compiled into one combined answer. Afterwards the presentation of these combined answers can be handled exactly the same as the Single or Multiple Choice class.
Another way is to use the Single or Multiple Choice class representation as starting point by using only the questions answer for it and provide the additional attributes using another object (e.g. list box). This object would reveal the additional information alongside the Single or Multiple Choice class representation.

4.4 Spreadsheet Question Classes analysis

Every SSQC requires an appropriate visual analysis in the dashboard of the external analytical application. A closer look reveals that there are synergies between the representation and the data of the different SSQCs. In other words, data from one class can be represented in a meaningful way using different representation forms from other classes. Figure 7 depicts the way different data can be displayed in various representations.
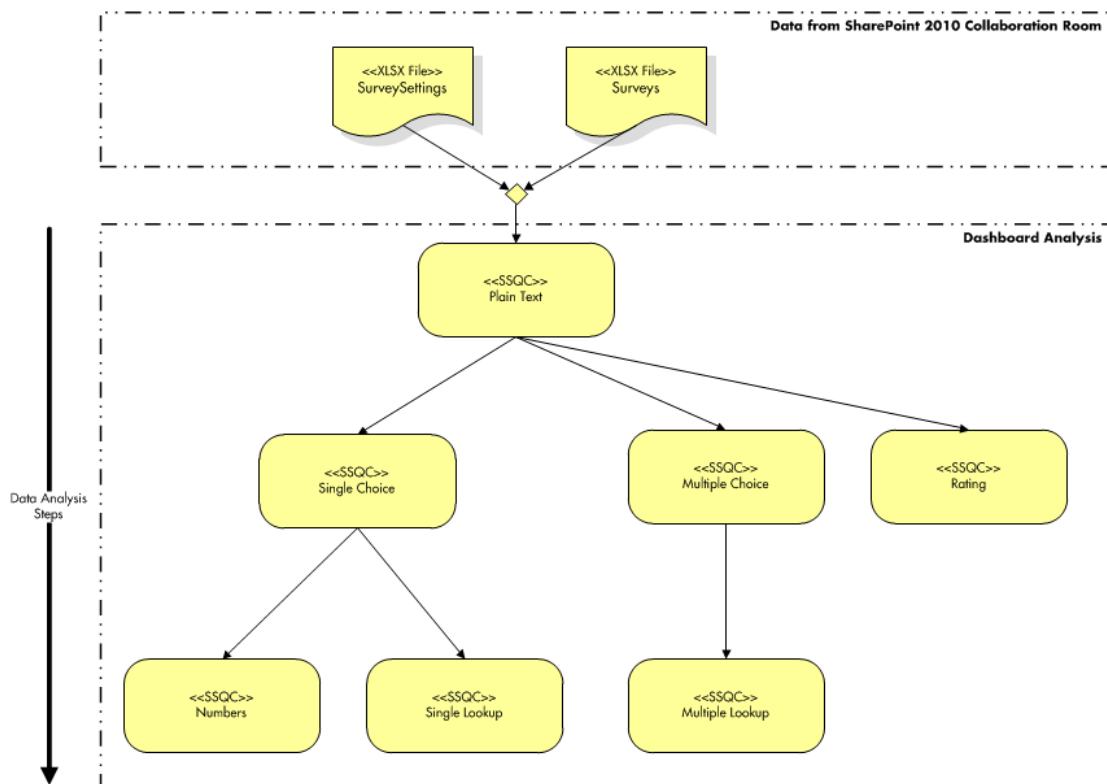
Figure 7: Spreadsheet Question Classes representation dependencies

All data collected and imported in the application can be visualized as a simple Plain Text analysis. Continuing from the general representation, data visualization methods may further specialize in more suitable and meaningful ways for other classes.

For example the Numbers class can use the Numbers, the Single Choice and the Plain Text representation. In this case all three different analysis methods are adding value to the evaluation of this question.

5 Evaluation of analysis tools

To evaluate a SharePoint survey it is not necessary to use other tools than SharePoint provides itself. But SharePoint limits the way data can be analyzed by presenting just a few static reports, which is the major reason for the evaluation of SharePoint surveys using an external analysis tool.

Before developing an application for the analysis of survey data, a suitable tool needs to be identified. The following chapter deals with the preconditions and the process of selecting the tool as well as the reason for choosing it.

5.1 The goal

The main goal was to develop an application which can evaluate surveys better than SharePoint and of course be able to evaluate all surveys available in the YPP (Young Professional Potentials) Collaborations room.

A collaboration space provides a place to collaborate for people, which enables personal interaction and information exchange. (Richardson 2011) A collaboration room is just another word for collaboration space. In the case of SharePoint this very same idea is transferred to a digital room with the extra benefit of storing additional information on the platform.

In this digital collaboration room the following surveys are available as test data:
- mySite - Getting Started
- mySite - Part 1 - Overview
- mySite - Part 2 - My Profile
- My Site - Part 3 – Tags, Notes and Bookmarks
- My Site - Part 4 – Colleagues, Memberships & Organization Chart
- Survey - Part X - Conclude

To provide sufficient visualization of the survey data, the following SSQC require consideration in the application:

- Plain Text
- Single Choice
- Multiple Choices
- Rating

That is why the application will primarily be designed to represent these four SSQC.

5.1.1 Selection Criteria for a tool

The following aspects need to be considered when selecting a suitable tool:

- Easy to use
- Fast and flexible analysis
- Use the spreadsheet export function in Excel in order to facilitate for the survey analyst to provide data to the application.
- Application can be filled with data without touching any kind of script, programming or visualization components. Therefore it should suffice to provide new spreadsheet files in order to analyze complete different surveys.

5.2 Comparison of analysis tools

There are various ways to visualize data. As the analysis tools should be provided to non-specialized users there are certain criteria to be taken into account. The first criterion demands a tool that is easy to understand and use. That means a tool that is familiar to the user or is just simple enough to interact with. The second criterion is to offer fast and flexible analysis capabilities.

To fulfill both criteria the selection has been narrowed down to the following BI tools:

- Spotfire from TIBCO
- QlikView from QlikTech
- Tableau Desktop from Tableau Software

- SQL Server Analysis Services (SSAS) Cubes from Microsoft

All these tools are capable of real-time data analysis by using in memory computing, which is very effective in delivering a fast response time. (Schaffner, et al. 2011) This is also the reason why these tools are on the shortlist.

Continuing there are two figures that compare these four tools. Figure 8 shows the comparison between data visualization tools done in 2010; bold markings define the leading tool in this category.

Criteria	Spotfire	Qlikview	Tableau	MS BI Stack	Comment
Business Criteria	=======	=======	=======	=======	**Speed, Scalability, Price**
Implementation Speed	Good	**High**	Good	Average	*Qlikview is fastest to implement*
Scalability	**Unlimited**	**Limited by RAM**	Very Good	Good	Need the expert in scalable SaaS
Pricing	High	Above Average	High	**Average**	*Microsoft is the price leader*
Licensing/support cost	High	High	High	**Average**	Smart Client is the best way to save
Enterprise Readiness	**Excellent**	Good for SMB	Good for SMB	**Excellent**	*Partners are the key to SMB market*
Long-term viability	Good	1 product	Average	**Excellent**	Microsoft are 35+ years in business
Mindshare	Analytics Market	**Growing fast**	**Growing fast**	3rd attempt to win BI	*Qlikview is a DV Leader, Successful IPO*
Technical Criteria	=======	=======	=======	=======	**Drilldown, Analytics, UI**
Clients for End Users	**ZFC, Spotfire Silver**	RIA, ZFC, **Mobile**	Windows, ZFC	**Excel, .NET**	*Free Qlikview Personal Edition is a big plus*
Interactive Visualization	Very Good	**Excellent**	Very Good	As good as Excel	Most users value Visualization over Modeling
Data Integration	Good	Good	**Excellent**	Good	Need for Data Integration expert
Visual Drill-Down	Very Good	**Excellent**	Good	Average	*Qlikview is fastest thanks to in-memory database*
Dashboard Support	Very Good	**Excellent**	Good	Below Average	Spotfire and Qlikview are best for Dashboards
Integration with GIS	**Excellent**	Good	Good	Average	Spotfire has the best GIS integration
Modeling and Analytics	**Excellent**	Weak	**Excellent OLAP**	Good with SSAS	*Spotfire is the best, Excel is the most popular*
UI & set of Visual Controls	Very Good	**Best**	Very Good	Good	Need for UI expert to integrate DV components
Development Environment	**Rich API, S+**	Scripting, Rich API	Average	**Excellent**	*Tableau requires less consulting than competitors*
64-bit In-Memory Columnar DB	Very Good	**Excellent**	In-memory Data Engine	Very Good	64-bit RAM allows huge datasets in memory

Figure 8: Data Visualization tools (Pandre, Comparison | Data Visualization 2010)

Figure 9 shows the updated version of the comparison in Figure 8, done in 2011. The numbers here represent how well a tool is fulfilling this task. The range is defined by 1 as the lowest and 9 as the highest. The numbers in red are highlighting the leading tool in this category.

	My Subjective Assessments			
Business Criteria	**Spotfire**	**Qlikview**	**Tableau**	**Microsoft**
Time to implement	6	9	6	2
Scalability	8	6	5	9
Price for Developer	5	4	6	8
Server License/user	5	3	5	5
Support fees / year	5	5	5	6
SaaS Platform	9	8	9	6
Overall Cost	4	4	5	2
Enterprise Ready	9	7	5	9
Long-term viability	8	7	6	9
Mindshare	5	6	7	4
Big Data Support	7	4	6	8
Partner Network	4	8	2	9
	My Subjective Comparison			
Visualization Criteria	**Spotfire**	**Qlikview**	**Tableau**	**Microsoft**
Data Interactivity	8	9	9	4
Visual Drilldown	7	9	8	2
Offline Viewer	8	7	9	6
Analyst's Desktop	8	9	9	4
Dashboard Support	8	9	8	5
Web Client	9	8	9	4
Mobile Clients	7	9	8	2
Visual Controls	8	8	8	5
UI Interactivity	8	8	8	4
	My Subjective Estimates			
Technical Criteria	**Spotfire**	**Qlikview**	**Tableau**	**Microsoft**
Data Integration	7	7	9	8
Development	8	7	5	9
64-bit in-memory DB	8	9	7	9
64-bit Desktop Client	7	8	1	9
Integration with GIS	8	6	9	5
Modeling, Analytics	9	4	5	6
Data Mining	7	2	3	8
Multidimensional Cubes	2	3	7	9
VertiPaq Support	1	1	7	9
PowerPivot Support	1	1	7	9

Figure 9: Data Visualization tools update (Pandre 2011)

5.3 Reasons for QlikView

All tools can handle one or more Excel sheets as input data. Considering the visualization aspects, the dashboard and the Graphical User Interface we reach a"…3-way tie, it … depends on personal preferences, but QlikView is more easy to use than others". (Pandre 2010) Out of the three tools left (Spotfire, QlikView and Tableau) QlikView was the only one available for a case study within the cooperative company and was therefore selected as analyzing tool.

6 QlikView SPSurvey Dashboard solution

This chapter is about the QlikView *SPSurvey Dashboard* solution, which is responsible for loading a predefined number of surveys into the application and will provide it afterwards for visualization and analyses. The data in general is collected through the standard SharePoint function *Export to spreadsheet* and some additional manually added information.

6.1 Application architecture

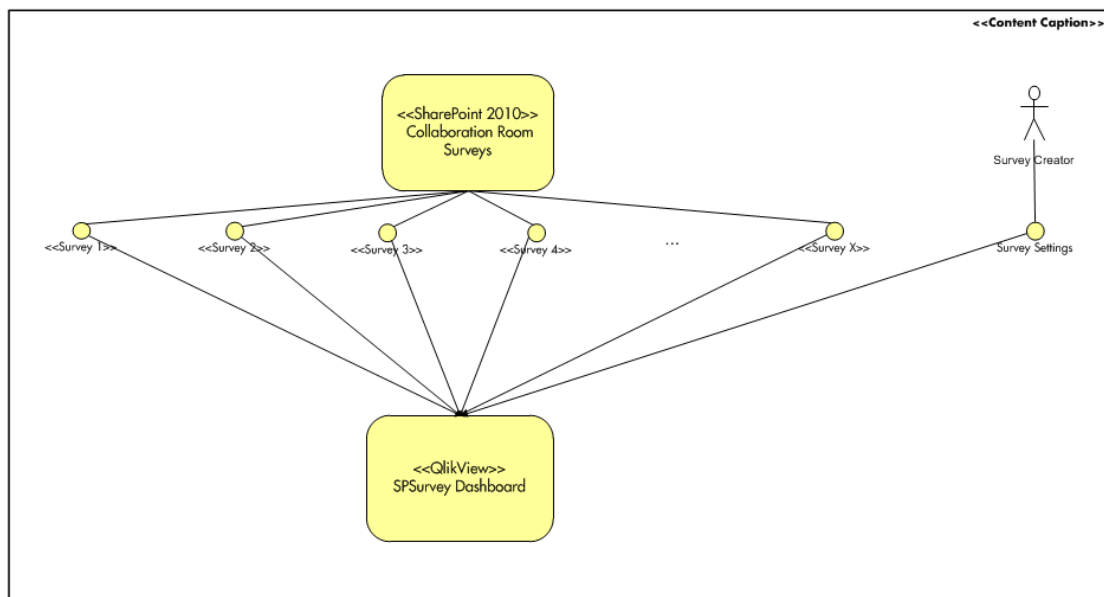The following diagram (Figure 10) depicts the application architecture of the QlikView application SPSurvey Dashboard.

Figure 10: Application architecture - SPSurvey Dashboard

Fehler! Verweisquelle konnte nicht gefunden werden. shows the interfaces connected to the QlikView application, and explains from where and what kind of data is flowing into the application.

The <<Content Caption>> entitles the content related part of the data. Further the <<Share Point 2010>> rectangle represents the SharePoint 2010 server with surveys from the Collaboration room as a single entity. The Survey Creator represents a real person generating data from the Collaboration Room. The circles are interfaces representing the data flow of the surveys into the QlikView application. An arbitrary number of surveys can be loaded into the application. The QlikView analytical application itself is identifiable by this notation: <<QlikView>>.

6.1.1 Description of Interfaces

The following Table 1 lists a detailed description of all interfaces represented within the application architecture.

Interface	Source System	Target System	Reload Frequency	Description
Survey 1	Collaboration Room Surveys	SPSurvey Dashboard	manually	Spreadsheet output of survey 1
Survey 2	Collaboration Room Surveys	SPSurvey Dashboard	manually	Spreadsheet output of survey 2
Survey 3	Collaboration Room Surveys	SPSurvey Dashboard	manually	Spreadsheet output of survey 3
Survey 4	Collaboration Room Surveys	SPSurvey Dashboard	manually	Spreadsheet output of survey 4
…	…	…	…	…
Survey X	Collaboration Room Surveys	SPSurvey Dashboard	manually	Spreadsheet output of survey X
Survey Settings	Survey Creator	SPSurvey Dashboard	manually	Creates all additional information which is not covered by the spreadsheet output from SharePoint

Table 1: Description of Interface

6.2 Software-Architecture

This chapter deals with the reason and the way the QlikView application was built. It shows the structure, the model and the scripts of the application and describes

furthermore what the input data for the application consists of and it's ensuing processing within the application.

The last section is about data representation, and explains how each class needs to be visualized in order to analyze the data in an efficient and comprehensive way.

6.2.1 Software-Architecture-Diagram

Figure 11 illustrates the software architecture of the QlikView application SPSurvey Dashboard. It describes, in more detail than the application architecture, where the input data for the QlikView application is coming from.
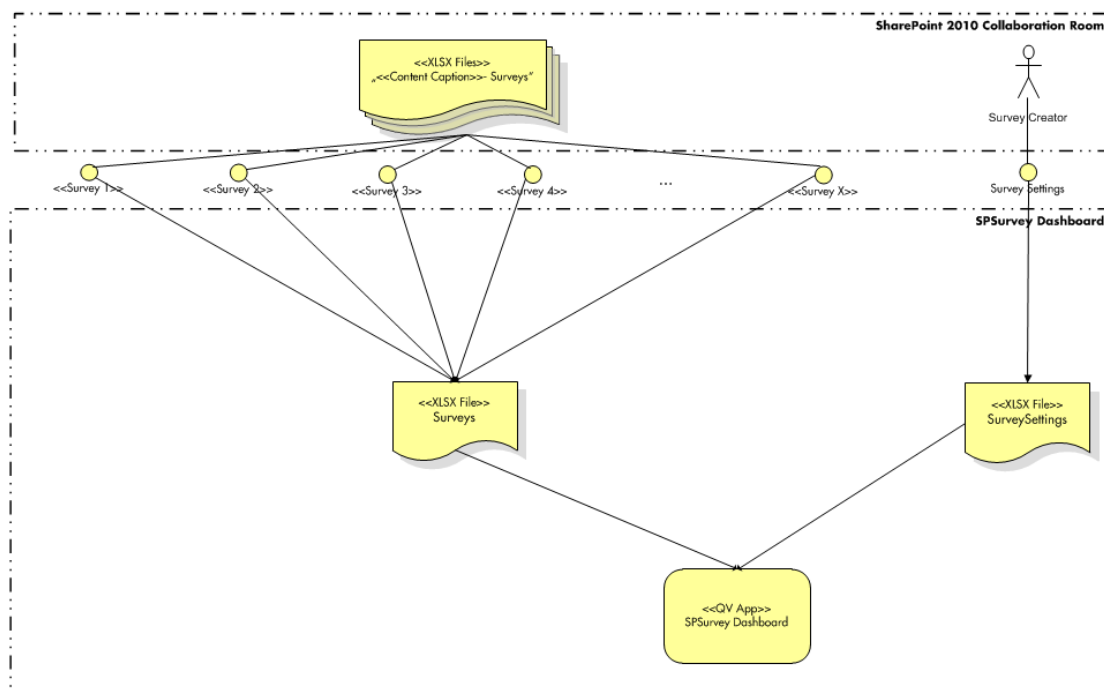

Figure 11: Software architecture - SPSurvey Dashboard

All surveys form the Collaboration room are exported via spreadsheets in SharePoint and put together into one single Excel file, surveys.xlsx, where each survey is represented as one tab (Excel worksheet) in the file.

Additionally there is a second file, SurveySetting.xlsx, containing all the Meta data missing from the data export.

Both files together are used as input for the QlikView application SPSurvey Dashboard.

6.2.2 Surveys

The Excel file surveys.xlsx holds all the exported spreadsheet data. The structure of the file is untouched, which means it is still the same data as was generated by SharePoint.

Table 2 shows the typical structure of a survey export:

Created By	1st Question	2nd Question	Item Type	Path
User 1	Answer 1	Answer 1;#Answer 2	Element	sites/ypp2011/Lists/SurveyName

Table 2: Example of the survey.xlsx

- *Created By*: The user who has given the answer
- *1st, 2nd, and so on Question*: This is the column for the question. The title row has always the same name as the question in SharePoint. For example *1. How much experience do you have with SharePoint?*
- *Item Type*: The content of the file always stays the same, namely *Element*. Hence this column is of no use for analyzing the survey.
- *Path*: Indicates the location of the survey on the SharePoint server, which is not needed for the task and therefore will not be evaluated during the survey.

Looking only at the data the type of question cannot be deducted.

6.2.3 SurveySetting

The file SurveySetting.xlsx contains all Meta information for the survey and also accompanying the surveys. All information in this file is entered manually by the Survey Creator, which is the person with access to all information and settings for each survey to be analyzed in QlikView.

The file is divided into several parts which are explained separately in the following subchapters.

6.2.3.1 Users

The *User* worksheet links all the users who are taking part in the survey with the department they are from.

The structure is shown in Table 3:

User	Department

Table 3: SurveySetting – Users

- *User*: The name of the user used by SharePoint to identify the his answers. The name has to be written exactly the same as in the survey export column *Created By*.
- *Department*: The name of the department or group the user is a part of.

The column name *Department* was selected because it represents the membership of the employee within the company and it fit for the surveys of the YPP Collaboration room. Furthermore it is possible to create other sorts of groups, like *Team Red*, Team Blue", and so forth if it is useful.

Knowing the connection between the user and the department facilitates the preservation of user anonymity and also enables a more powerful analysis of the data.

6.2.3.2 Departments

This worksheet pinpoints the color for each department, shown in the Table 4.

Department	R	G	B

Table 4: SurveySetting - Departments

- *Department*: The name of the department
- *R*: The red proportion of the RGB (red, green and blue) color space.
- *G*: The green proportion of the RGB color space.
- *B*: The blue proportion of the RGB color space.

Storing the color information of the department externally opens the possibility of individual color settings for the analyst and also provides a consistent color setting within the application for all charts.

6.2.3.3 SurveyOverview

All surveys accessible through the SPSurvey Dashboard will be listed using the structure shown in Table 5.

SurveyNr	SurveyName	SurveyTabName

Table 5: SurveySetting – SurveyOverview

- *SurveyNr*: The unique survey key. It is needed to establish a connection between questions, answers and ratings.
- *SurveyName*: Name shown in the QlikView application for that survey.
- *SurveyTabName*: Name of the worksheet in the survey.xlsx file.

This SurveyOverview list serves as basis for loading surveys into the application. If a survey is not listed here it will not be loaded into the SPSurvey Dashboard.

6.2.3.4 QuestionOverview

This QuestionsOverview list contains all questions of the surveys using the structure shown Table 6.

QuestionName	QuestionTitle	QuestionSurveyNr	SurveyNr	QuestionType

Table 6: SurveySetting – QuestionOverview

- *QuestionName*: This is the same name as was exported from SharePoint (column title)
- *QuestionTitle*: Display name of the question in the QlikView application.
- *QuestionSurveyNr*: The unique question key to establish a connection between answers and ratings.
- *SurveyNr*: The unique key of the survey within this Excel file.
- *Question Type*: The SSQC of the question. Defines what the question is referring to.

6.2.3.5 AnswersOverview

All answers to questions of the Single or Multi Choice sort are stored here. The main purpose of this sheet is to define the sorting order of the answer to reflect the way they were displayed to the user, shown in Table 7.

AnswerName	QuestionSurveyNr	SurveyNr	SortNr

Table 7: AnswersOverview

- *AnswerName*: The complete answer.
- *QuestionSurveyNr*: The unique key of the question in the SurveySettings document.
- *SurveyNr*: The unique key of the survey within the Excel file.
- *SortNr*: Defines in which order the answers should be displayed. Sorting is done with numbers.

6.2.3.6 RatingOverview

The last sheet is responsible for storing all range information of the rating questions. This information classifies the answer of the user in more detail. Each sub-question of one super rating class is stored separately in the sheet, although all values are the same. The column names are listed in Table 8.

QuestionSurveyNr	SurveyNr	RangeLow	RangeMiddle	RangeHigh	NumberRange	NAOption

Table 8: RatingOverview

- *QuestionSurveyNr* : The unique question key of this file.
- *SurveyNr*: The unique key of the survey within this Excel file.
- *RangeLow*: Defines a textual expression of the lowest value of the rating.
- *RangeMiddle*: Defines a textual expression of the middle value of the rating.
- *RangeHight*: Defines a textual expression of the highest value of the rating.
- *NumberRange*: Defines how many steps a range question consists of. For example a NumberRange five means that the value one is the lowest and five is the highest.
- *NAOption*: Defines the text of the N/A option. This field is optional. It has to be empty when the N/A option in SharePoint is not selected.

6.2.4 SPSurvey Dashboard

The SPSurvey Dashboard application loads data of both Excel files and generates one combined data model, which will be described later on. This model then provides the data for an exact analysis of the surveys, which is also done in this application.

The whole process of loading the data and transforming it is described by a QlikView script done in the QlikView Script Editor.

6.2.4.1 Dimensions

At first the dimension has to be built, starting by loading the user table from the SurveySettings.xlsx file. Code Sample 1 shows the responsible script part.

```
Users:
LOAD Hash128(User) as %User,
     Department as %Department
FROM
LocalData\SurveySettings.xlsx
(ooxml, embedded labels, table is Users);


//Generate Fake Users
Concatenate (Users)
LOAD Hash128('TEST_' & %Department) as %User,
     %Department
Resident Department;
```

Code Sample 1: Load user data

Code Sample 2 illustrates one of the data security measurements. The name of the user is stored as a hash value in the data model, whereas the original name will not be stored there. So the whole analysis is anonymous, only the users are connected to a group - the department.

```
Department:
LOAD Department as %Department,
       R,
       G,
       B
FROM
LocalData\SurveySettings.xlsx
(ooxml, embedded labels, table is [Departments]);


//Generate Fake Users
Concatenate (Users)
LOAD Hash128('TEST_' & %Department) as %User,
     %Department
Resident Department;
```

The department table is loaded with the color set for each department. Additionally some fake data is generated, which is necessary to ensure that all departments are displayed in the charts, even if no user from the department took part in the survey.

```
SurveyOverview:
LOAD SurveyNr as %SurveyNr,
       SurveyTabName as TabName,
       SurveyName as Name
FROM
LocalData\SurveySettings.xlsx
(ooxml, embedded labels, table is [SurveyOverview]);
```

Code Sample 3: Load SurveyOverview data

Code Sample 3 depicts how all SurveyOverview data is loaded into the application.

```
QuestionOverview:
LOAD autonumber(SurveyNr & '/' & QuestionSurveyNr) as %QuestionKey,
       SurveyNr as %SurveyNr,
       RowNo() as RowNo,
       QuestionSurveyNr,
       QuestionName as Name,
       QuestionTitle as Title,
       QuestionType as Type
FROM
LocalData\SurveySettings.xlsx
(ooxml, embedded labels, table is [QuestionOverview]);
```

Code Sample 4: QuestionOverview information

Code Sample 4 on the other hand loads the QuestionOverview data into the application. A special %QuestionKey is generated to define a primary key for this

table. The RowNo will be used later on to determine how many questions the whole analysis will cover.

```
AnswersOverview:
LOAD autonumber(SurveyNr & '/' & QuestionSurveyNr & '/' & AnswerName) as
%AnswerKey,
        SurveyNr as SurveyNr,
        QuestionSurveyNr as QuestionNr,
        AnswerName as Answer,
        SortNr as SortNr
FROM
LocalData\SurveySettings.xlsx
(ooxml, embedded labels, table is [AnswersOverview]);
```

Code Sample 5: AnswersOverview information

Code Sample 5 loads the AnswersOverview into the application. The primary key %AnswerKey is generated using the information from the excel sheet.

```
RatingOverview:
LOAD autonumber(SurveyNr & '/' & QuestionSurveyNr) as %QuestionKey,
        RangeLow,
        RangeMiddle,
        RangeHigh,
        NumberRange,
        NAOption
FROM
LocalData\SurveySettings.xlsx
(ooxml, embedded labels, table is [RatingOverview]);
```

Code Sample 6: RatingOverview information

The Code Sample 6 shows the load of the RatingOverview table with the primary key %QuestionKey.

6.2.4.2 Fact Table

The centerpiece of the application is the *Survey* table, which contains all answers given by the users during the survey. Hence the data will be loaded directly from the survey.xlsx.

The following fields are stored within the table:

- *%QuestionKey*: A combination of question and survey key to link the Question and the Rating overview table.
- *%User*: Hashed user to get information about their department.
- *%AnswerKey*: The primary key of this table.
- *Survey.AnswerName*: The answers by the user during the survey, loaded from the file survey.xlsx.
- *Survey.Valid*: Defines the validity. 1 is valid and equal or less than vSurveyValidNumber is not.

The variable *vSurveyValidNumber* is 0.0000001. It is important to have some kind of true value unlike zero, which is also very small. Then, if the column *Survey.Valid* is summed up in the dashboard the invalid values will not have an impact on the number of the valid due to rounding.

The starting point to create the fact table is the QuestionOverview table. All data related to the question will be loaded and generated considering each entry,shown in Code Sample 7.

```
LET vQuestionTotal = Peek('QuestionOverview.RowNo', -1, 'QuestionOverview');
...
for vQuestionCount = 0 To ($(vQuestionTotal) - 1)
      ...
NEXT
```

Code Sample 7: Fact table – handle each question separately

Afterwards the data processing is divided into four parts, depending on the SSQC of the Question.

- MultipleChoice
- SingleChoice
- Plain Text
- Rating

In the previous chapter a few more SSQC classes were introduced but only four of them were used in the YPP Surveys and subsequently implemented in the application.

The handling of these four SSQC are explained in detail below.

6.2.4.2.1 Multiple Choice

Basically all data covering a particular multiple choice question will be loaded in a temporary table called Answers. Code Sample 8 illustrates the approach for multiple choice answers.

```
Answers:
LOAD autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr) & '/' &
[$(vQuestionName)]) as %AnswerKey,
                   [Created By] as %User,
                   [$(vQuestionName)] as Survey.AnswerName,
                   RowNo() as Answers.RowNo
FROM
LocalData\surveys.xlsx
(ooxml, embedded labels, table is [$(vSurveyLocation)]);
```

Code Sample 8: Fact table – Multiple Choice Answers

Due to the way the values are stored by SharePoint for this question it is necessary to break up the field Survey.AnswerName even more. More detailed information on SSQC Multiple Choice is provided in chapter 4.3.1.3. The fragmentation of the Survey.AnswerName is performed using the following Code Sample 9:

```
LET vAnswersTotal = Peek('Answers.RowNo', -1, 'Answers');
```

```
for vAnswersCount = 0 to $(vAnswersTotal)-1


     LET vAnswerKey = Peek('%AnswerKey', $(vAnswersCount), 'Answers');
     LET vUser = Peek('%User', $(vAnswersCount), 'Answers');
     LET vServeyAnswers = Peek('Survey.AnswerName', $(vAnswersCount),
'Answers');
     LET vAnswersUserTotal = SubStringCount('$(vServeyAnswers)', ';#') + 1;


     for vAnswersUserCount = 1 to $(vAnswersUserTotal)
        ...
     NEXT
NEXT


DROP Table Answers;
```

Code Sample 9: Fact table – Multiple Choice subdivide Survey.AnswerName

If the user gave more than one answer to one question, each answer is noted separately. For example the value of one answer is *Edit my profile;#Add or remove colleagues;#Use function* the loop will iterate three time to generate three arguments for this user:

- Edit my profile
- Add or remove colleagues
- Use function

The following Code Sample 10 is generating these rows and links the result with the already existing one.

```
if len(KeepChar('$(vServeyAnswers)',
'abcdefghijklmnopqrstuvwxyzäöüABCDEFGHIJKLMNOPQRSTUVWXYZÄÖÜ1234567890?\()/&%$
§"!{}()')) > 0
THEN


     Let vAnswer = subfield('$(vServeyAnswers)',';#',$(vAnswersUserCount));
ELSE
```

```
        Let vAnswer = 'N/A';
ENDIF


Concatenate (Survey)
LOAD autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr) & '/' &
'$(vAnswer)') as %AnswerKey,
        autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)) as %QuestionKey,
        Hash128(%User) as %User,
        Survey.AnswerName,
        1 as Survey.Valid
INLINE [
        %AnswerKey, %User, Survey.AnswerName
        $(vAnswerKey), $(vUser),  $(vAnswer)
                                        ];
```

Code Sample 10: Fact table – Multiple Choice answers concatenate

Ultimately it is necessary to create fake data, shown in Code Sample 11, to ensure
that every answer is displayed in the dashboard charts. If this data is not generated,
all answers which were not selected by any user during the survey will be omitted and
invisible to analysis.

```
LET vFirstDepartment = Peek('%Department', 0, 'Department');


Concatenate (Survey)
LOAD autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr) & '/' &
AnswersOverview.Answer) as %AnswerKey,
        autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)) as %QuestionKey,
        Hash128('TEST_' & '$(vFirstDepartment)') as %User,
        AnswersOverview.Answer as Survey.AnswerName,
        $(vSurveyValidNumber) as Survey.Valid
Resident AnswersOverview
WHERE       AnswersOverview.SurveyNr = '$(vSurveyNr)' and
            AnswersOverview.QuestionNr = '$(vQuestionSurveyNr)';
```

Code Sample 11: Fact table – Multiple Choice Fake Data

6.2.4.2.2 Single Choice

The SSQC Single Choice simply adds all data to the Survey table. The Code Sample 12 is less complex because every user can give exactly one answer.

```
Concatenate (Survey)
LOAD autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr) & '/' &
[$(vQuestionName)]) as %AnswerKey,
        autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)) as %QuestionKey,
        Hash128([Created By]) as %User,
        [$(vQuestionName)] as Survey.AnswerName,
        1 as Survey.Valid
FROM
LocalData\surveys.xlsx
(ooxml, embedded labels, table is [$(vSurveyLocation)]);
```

Code Sample 12: Fact table – Single Choice Load

Like the Multiple Choice class, this class also generates fake data to improve the readability of data in the dashboard. The code is exactly the same as before in Code Sample 11 for the Multiple Choice class.

6.2.4.2.3 Plain Text

The variable vQuestionType represents the class of the question. If the value differs from *SingleChoice* or *MultipleChoice* the following loading script, Code Sample 13, is used for the last classes *PlainText* or *Rating*:

```
Concatenate (Survey)
LOAD autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr) & '/' &
[$(vQuestionName)]) as %AnswerKey,
        autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)) as %QuestionKey,
        Hash128([Created By]) as %User,
        if(len(KeepChar([$(vQuestionName)],
'abcdefghijklmnopqrstuvwxyzäöüABCDEFGHIJKLMNOPQRSTUVWXYZÄÖÜ1234567890?\()/&%$
§"!{}()')) > 0, [$(vQuestionName)], 'N/A') as Survey.AnswerName,
        if(len(KeepChar([$(vQuestionName)],
```

```
'abcdefghijklmnopqrstuvwxyzäöüABCDEFGHIJKLMNOPQRSTUVWXYZÄÖÜ1234567890?\()/&%$
§"!{}()')) > 0, 1, 0) as Survey.Valid
FROM
LocalData\surveys.xlsx
(ooxml, embedded labels, table is [$(vSurveyLocation)]);
```

Code Sample 13: Fact table – Plain Text Load

The load statement is similar to the others except for the statements behind the Survey.AnswerName and Survey.Valid. Sometimes SharePoint returns some unreadable ASCII characters instead of an empty field. Therefore it is necessary to know whether the user has entered something or not.

No fake data has to be generated for this class because no answer is known in advance anyway.

6.2.4.2.4 Rating

The Rating class connects directly to the load statement of the Plain Text class, but it is again necessary to add some fake data. This is done in two steps. The first step covers every kind of rating number, shown in Code Sample 14:

```
for vRatingCount = 1 to $(vNumberRange)


      Concatenate (Survey)
      LOAD autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr) & '/' &
'$(vRatingCount)') as %AnswerKey,
            autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)) as
%QuestionKey,
            Hash128('TEST_' & '$(vFirstDepartment)') as %User,
            '$(vRatingCount)' as Survey.AnswerName,
            $(vSurveyValidNumber) as Survey.Valid
      INLINE [
            %AnswerKey, %User, Survey.AnswerName
            1, 1,  1
                  ];
```

```
NEXT
```

Code Sample 14: Fact table –Load fake data for rating numbers

The INLINE load is just a fake load in order to be able to add a new row with data.

The second step, represented by Code Sample 15, generates a fake data set for the N/A value if it is available for this question.

```
LET vNAOption = lookup('RatingOverview.NAOption', '%QuestionKey',
autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)), 'RatingOverview');

IF len('$(vNAOption)') > 0 THEN

Concatenate (Survey)
LOAD autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr) & '/' &
'$(vNAOption)') as %AnswerKey,
        autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)) as %QuestionKey,
        Hash128('TEST_' & '$(vFirstDepartment)') as %User,
        '$(vNAOption)' as Survey.AnswerName,
        $(vSurveyValidNumber) as Survey.Valid
INLINE [
        %AnswerKey, %User, Survey.AnswerName
         1, 1,  1
                    ];

ENDIF
```

Code Sample 15: Fact table – Rating load fake data for N/A option

6.2.4.2.5 Additional Fake Data

Now all data from the survey.xlsx and the surveysettings.xlsx files is loaded into the application. Again some fake data has to be generated to ensure each department is

listed and will be displayed in the charts even if no department member has given an answer to the corresponding question. Code Sample 16 shows the implementation.

```
LET vAnswerKey = Peek('%AnswerKey', -1, 'Survey');
LET vAnswer = Peek('Survey.AnswerName', -1, 'Survey');


Concatenate (Survey)
LOAD '$(vAnswerKey)' as %AnswerKey,
        autonumber($(vSurveyNr) & '/' & $(vQuestionSurveyNr)) as %QuestionKey,
        Hash128('TEST_' & %Department) as %User,
        '$(vAnswer)' as Survey.AnswerName,
        $(vSurveyValidNumber) as Survey.Valid
Resident Department;
```

Code Sample 16: Fact table – Additional fake data

6.2.4.3 Data model

The output of the previous load statements results in the QlikView relational data model, which is based on a snowflake schema. (See Figure 12)
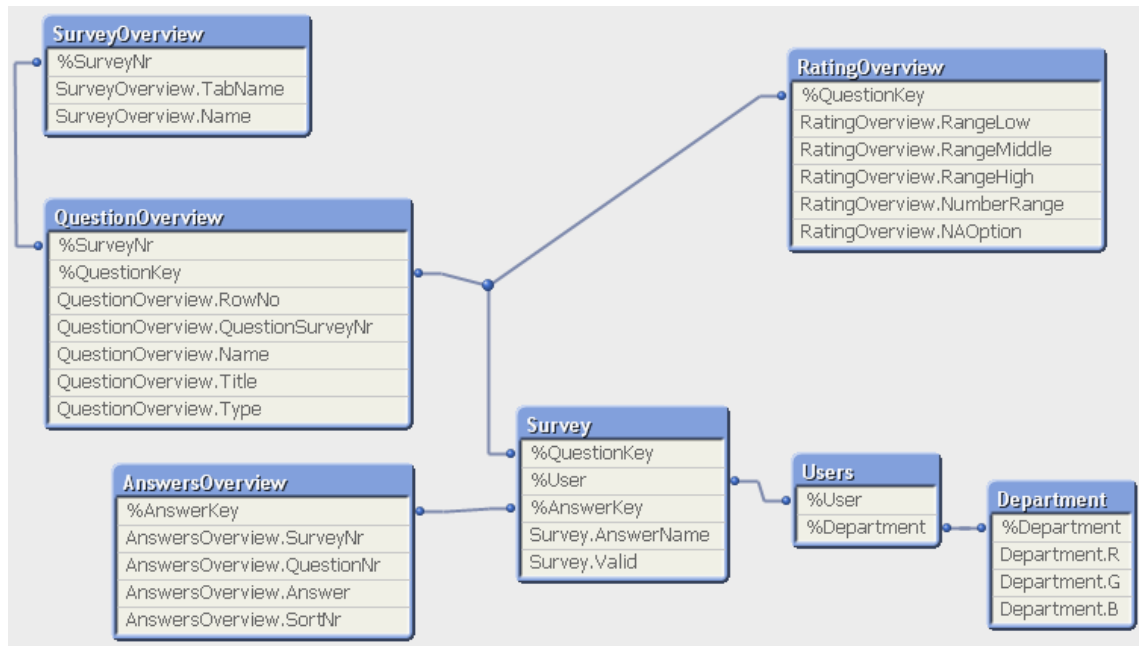
Figure 12: QlikView Datamodel - SPSurvey Dashboard

To adequately introduce the snowflake schema the star pattern has to be explained first. A star pattern is a relational data model defined by two types of tables. The first one is the fact table holding data which is not predefined and can change rapidly. In contrast the dimension tables contain information that describes the fact data and they change less often. In general a star schema is defined by one or more fact tables connected by dimension tables. That means all fact values are stored in a few or even only one table which is connected to all dimensions table. Therefore the dimension tables themselves are not connected to each other. (Richter 2001)

The snowflake schema however expands the star model by splitting the dimension tables and connecting them together, which led to the name giving form. Therefore the data can be normalized more. (Manhart 2008)

As to our model, the answers given by the users represent the fact data. Therefore they are stored in our fact table Survey. All other tables are dimension tables, which add more value to the fact data. The fact that some dimensions are not directly connected to the fact table is the reason why this model is a snowflake schema. For example the department table is linked via the user table to the Survey table.

The way QlikView tables are aligned in our model "… is 'associative' rather than relational, so lines between tables don't reflect foreign key relations. Instead, the lines mean that there is a like-named column in each of the related tables." (Lambert 2012)

6.2.4.4 Dashboard

The dashboard is the most important part in a QlikView application as it represents the data presented by the model to the user. In order to analyze the survey data every SSQC needs to be taken into account. Please remember, all data shown in the figures below are provided from surveys of the YPP Collaboration Room. Because the application initially was built to analyze this survey, the related data serves well to show the functionality and usage of the dashboard.

The following sub chapters describe the dashboard step by step, starting with an overview of the general objects and finishing with the objects needed, especially for specific SSQC.

6.2.4.4.1 Overview

While creating the application, the idea was to develop a compact application that offers as much information as possible at one glance. Therefore the application has been divided into just two different sheets (see below in Figure 13).

Figure 13: Dashboard - Sheets

The first one, *SPSurvey Dashboard – Home*, is an introductory sheet, shown in Figure 14. It informs the user about the application name and its functions.

Analyse SharePoint Surveys in QlikView

Version 1.0, release date 18.10.2013, reload date 07.11.2013 16:33:35

Figure 14: Dashboard - Application Logo & Title

The graphic visualization in Figure 14 illustrates the use of the application for the user and displays the applications name. Furthermore a few key data related to the application are provided:

- *Version Number*: The release number of the application.
- *Release Date*: The date of the official deployment, the release date.
- *Reload Date*: The timestamp of the last reload of the application. It also represents the data actuality.

The second tab, SPSurveys, shows all the representation possibilities of the prospective analyzed data. This screen, see Figure 15, is packed with information, charts, lists, tables, filters, text boxes, etc.
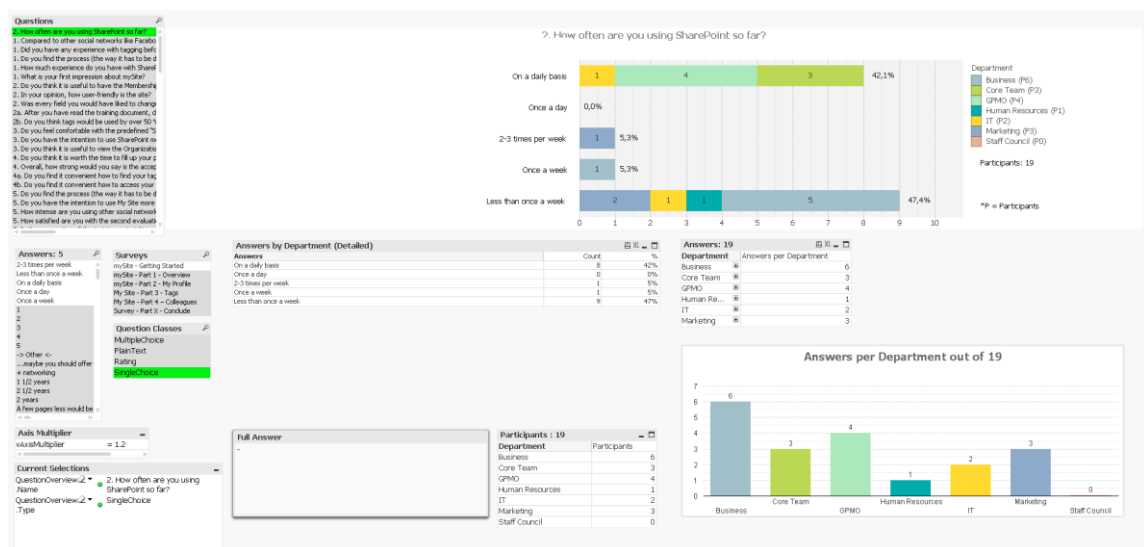
Figure 15: Dashboard – SPSurveys Overview

Figure 15 also represents an overview of the SPSurvey dashboard. All objects in this dashboard will be explained in the following chapter.

6.2.4.4.2 General objects

Apart from the objects especially designed to represent the data for each class, these general objects just present basic information about all the surveys. No direct analysis is possible with these but they help to narrow down the flood of information in order to focus on a slice of the data volume.
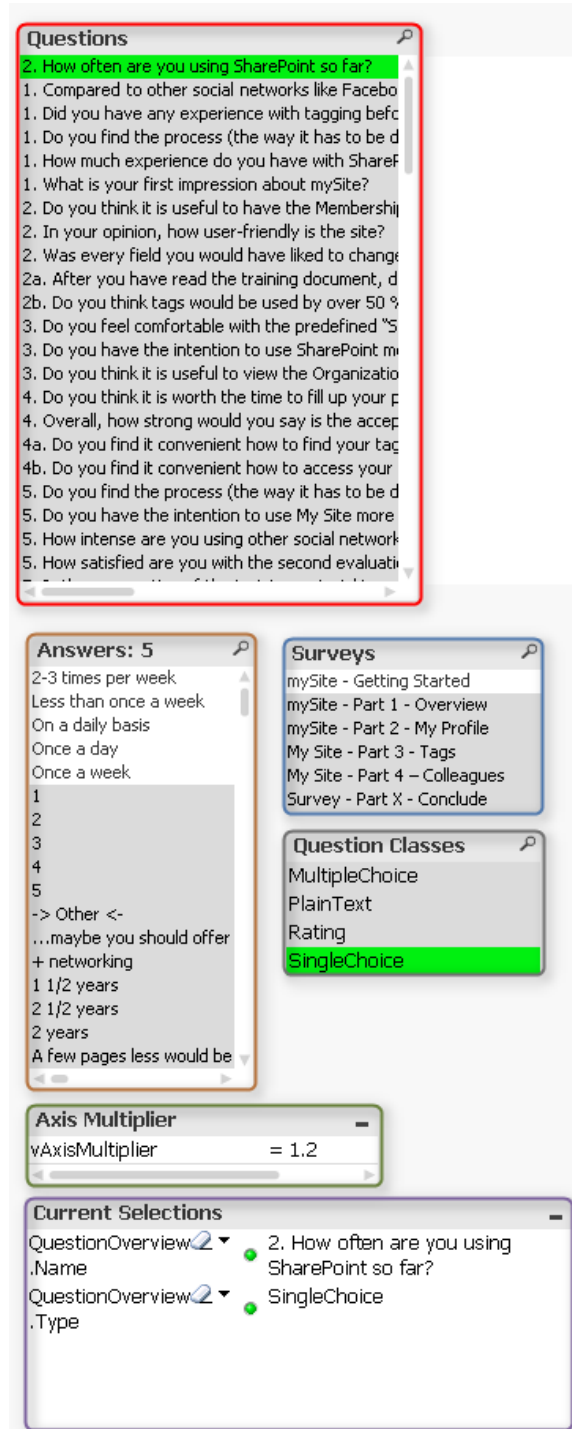
Figure 16: Dashboard – General objects

The general objects shown in Figure 16 are described in detail below.

- Questions: This list object lists all questions posted in all surveys.

- Answers: Lists all answers given by the users.
- Surveys: A list of all surveys that are presented in this loaded dataset.
- Question Classes: Contains all SSQC.
- Axis Multiplier: With this object the user can change the value of the variable *vAxisMultiplier*, which adjusts the length of the bars in order to squeeze or stretch the scale of the chart where the Ratings, Single Choice and Multiple Choice data is presented.
- Current Selections: Aids the user by showing his current selection in the application.

6.2.4.4.3 Plain text

The Plain Text analysis is the simplest form of looking into the data. The only purpose is to represent the data as it is because the data itself is self-written text and almost always unique. Therefore no information beyond the answers is available in advance.

As it was already explained for the SSQC Plain Text in the chapter 4.3.1.1, all objects introduced here can be used for analyzing other SSQC data. In other words all Plain Text objects support the analysis of all other SSQCs data. These objects are covered in Figure 17.
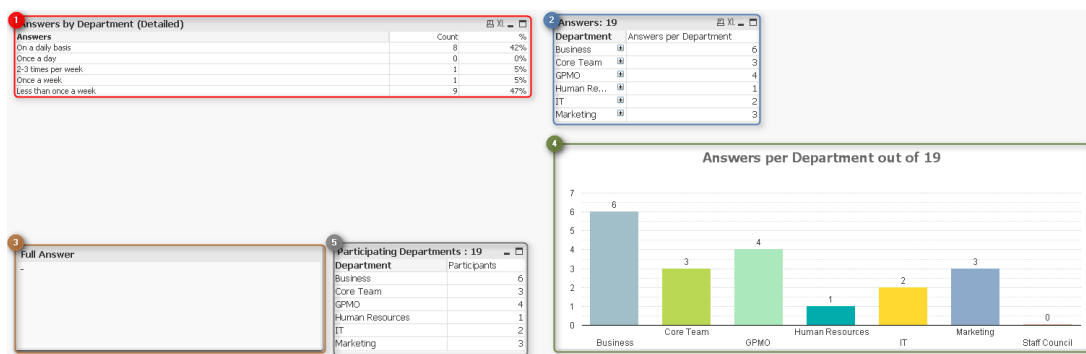

Figure 17: Dashboard – Plain text objects

The data used for this chapter and its subchapter was taken from the YPP survey *mySite – Getting Started* Question 14.

6.2.4.4.3.1 Answers by Department

Figure 18 shows a list box. This box contains all answers corresponding to the selected questions. Additionally the number of answers as well as the percentage of their relation to the other answers is displayed.

Answers by Department (Detailed)			
Answers		**Count**	**%**
corporate culture and the conservativeness of people within the bank are the reason for using means of…		1	8%
I wanted to help you to get the results.		1	8%
I was asked several times		1	8%
I was part of the original project team.		1	8%
interest		1	8%
it helps the trainer;		1	8%
notifications by email		1	8%
Pilot user		1	8%
Reminder from Markus Paulhart.		1	8%
to help a student		1	8%
to use My Site		1	8%
YPP Project		1	8%

Figure 18: Dashboard – Answers by Department

6.2.4.4.3.2 Answers

This Answers list-box is similar to the Answers by Department list-box. The main difference is the subdivision into Department. Here answers can be analyzed throughout each department. An example is shown in Figure 19.

Answers: 12		
Department	**Answer**	**Answers per Department**
Business	⊞	4
Core Team	⊞	1
GPMO	⊞	3
Human Re…	⊞	1
IT	⊞	1
Marketing	⊟ interest	1
	YPP Project	1

Figure 19: Dashboard – Answers

6.2.4.4.3.3 Full Answer

This text box will display the total content of an answer, if a single answer is selected within the application. This way even very long answers are readable, shown in Figure 20.
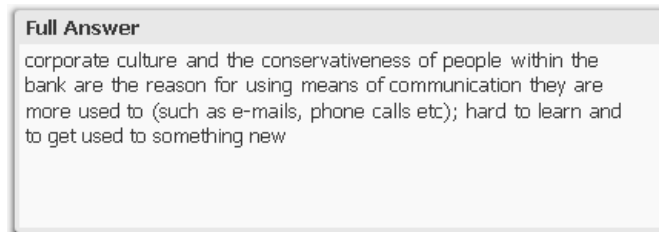

Full Answer

corporate culture and the conservativeness of people within the bank are the reason for using means of communication they are more used to (such as e-mails, phone calls etc); hard to learn and to get used to something new

Figure 20: Dashboard – Full Answer

6.2.4.4.3.4 Answers per Department out of

This chart shows all departments participating in the surveys, including the number of answers given by the department. Hence it always leads to a complete list of departments in this chart.

The additional information also highlights the department not participating in the survey. An example is shown in Figure 21: Dashboard – Answers per Department out of Figure 21.
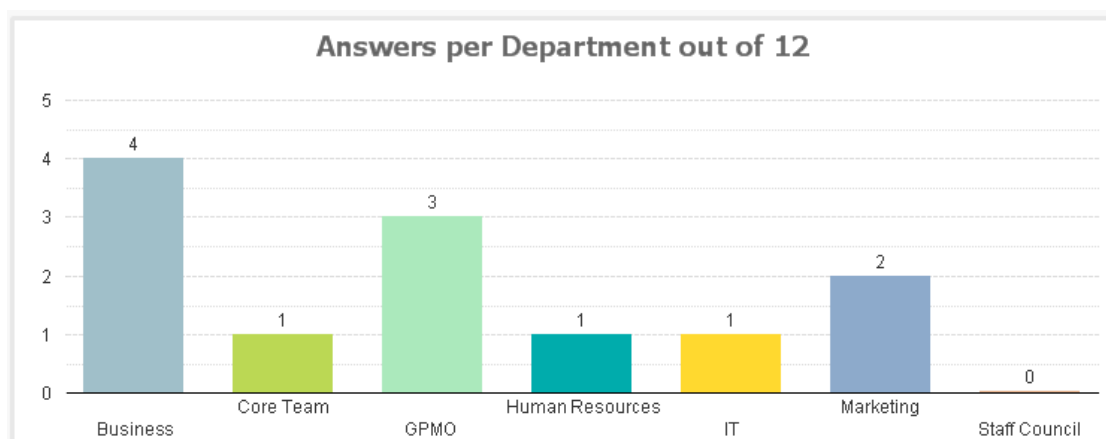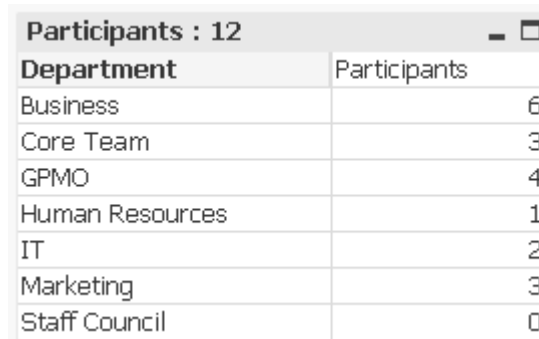
Figure 21: Dashboard – Answers per Department out of

The challenge in this chart was to show all departments even the ones not providing any answers to the question. Chapter 6.2.4.2.5 explains how this challenge is solved in the script part of the application.

6.2.4.4.3.5 *Participants*

This list shows all participants of the selected criteria. The number of the participants and the number of answers are not always corresponding. A good example for this would be data from the Multiple Choice SSQC.

It is important to have a separated list like in Figure 22.

Participants : 12	
Department	Participants
Business	6
Core Team	3
GPMO	4
Human Resources	1
IT	2
Marketing	3
Staff Council	0

Figure 22: Dashboard – Participants

6.2.4.4.4 Single Choice

While introducing the chart which represents Single Choice SSQC data this chapter also introduces the chart for the Multiple Choice and Rating SSQC. In fact the chart is the same for all of them but has a different appearance depending on the kind of data displayed.

This chapter describes the look and feel of the Single Choice SSQC data as shown in Figure 23.
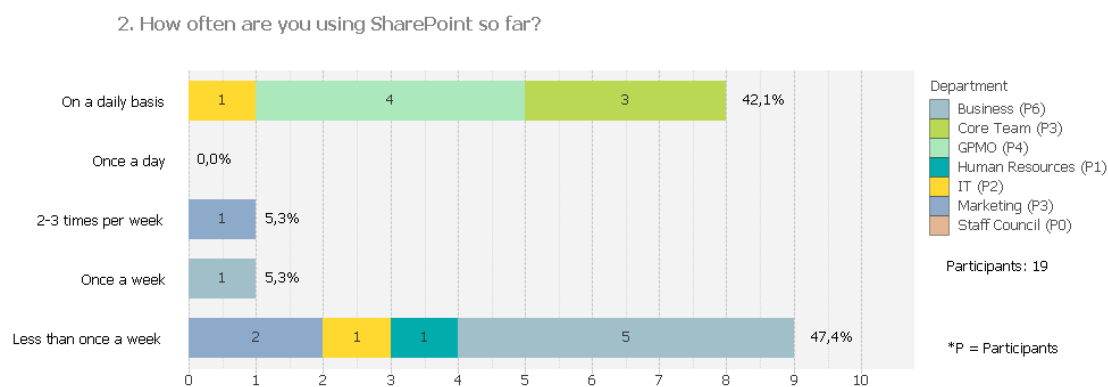
2. How often are you using SharePoint so far?

Figure 23: Dashboard – Single Choice

The view in Figure 23 illustrates all answers divided by the department also in relation to the total number. It displays all available answers for the particular question including answers which have not been chosen by anyone. The order of the answers is the same as for the participant in the survey.

To provide a more detailed view, the chart is divided into five visual areas depicted in Figure 24:

1. Title: Represents the question name.
2. Answer text: Shows all possible answers next to the corresponding graph of the graphic and numbers area.
3. Graphic and Numbers: Each answer bar represents the number of a unique answer grouped respectively to the department. Additionally this value is set in relation to all other answers displayed as a percentage value.
4. Agenda: The agenda field describes the colors in the graphic and numbers area an additionally provides information about the number of participants for this question with reference to the department. The number is highlighted using the letter *P* in front of it.
5. Additional Information: This block shows how many people in general were participated in this survey.
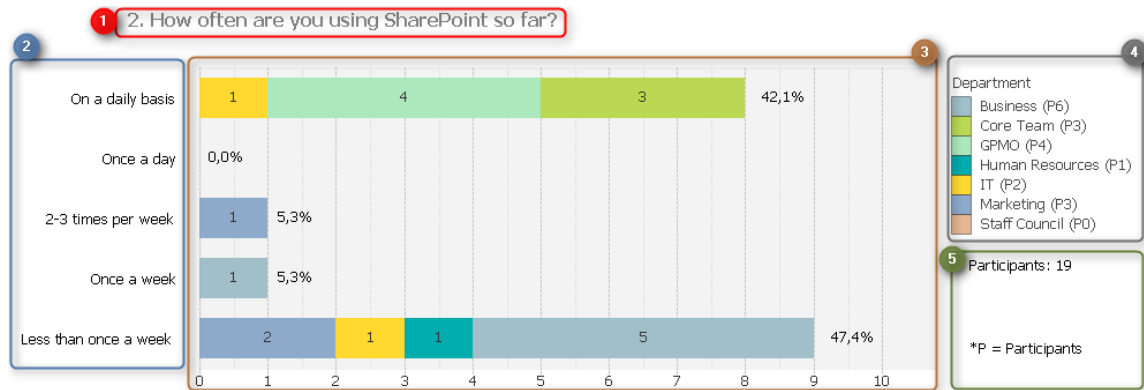
Figure 24: Dashboard – Single Choice visual areas

The data shown in Figure 23 and Figure 24 was taken from the YPP survey *mySite – Getting Started* Question 2.

6.2.4.4.5 Multiple Choice

As mentioned in the previous chapter, the chart representing the data for the SSQC Single Choice is the same as the one for the Multiple Choice. Therefore this chapter explains the differences in data display when a Multiple Choice question is analyzed.

The main different here is that multiple answers per participant can be displayed using only one chart. The approach is shown in Figure 25.
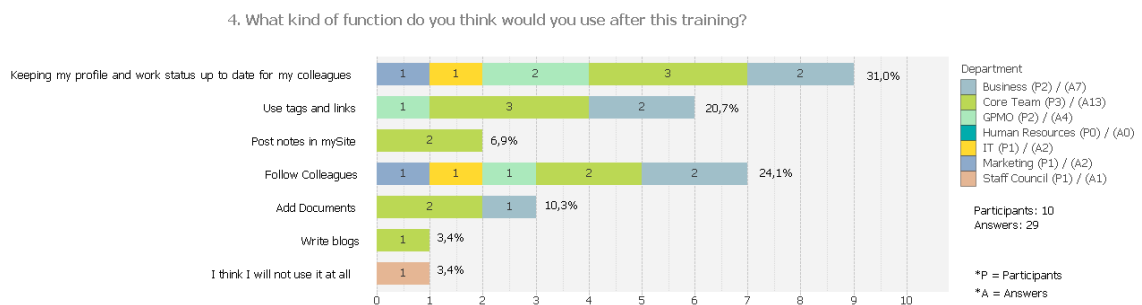
Figure 25: Dashboard – Multiple Choice

The chart is basically the same as the Single Choice representation, except that the answers are also counted.

The visualization area four and five, shown in Figure 26, will be described in more detail below:

4. Agenda: In addition to the previous agenda, the number of answers provided by the department is listed. The letter *A* in front of the number indicates this value.

5. Additional Information: Finally the total number of answers will be shown in this area.

Figure 26: Dashboard - Multiple Choice visual areas

All the other areas are the same as for the Single Choice chart.

6.2.4.4.6 Rating

This SSQC analysis chart is also generally displayed the same way as the Single Choice chart. The only difference is the representation of the answer text area. The basic chart display for rating data is shown in Figure 27.

4. How do you like the design of the My Site?_Colors
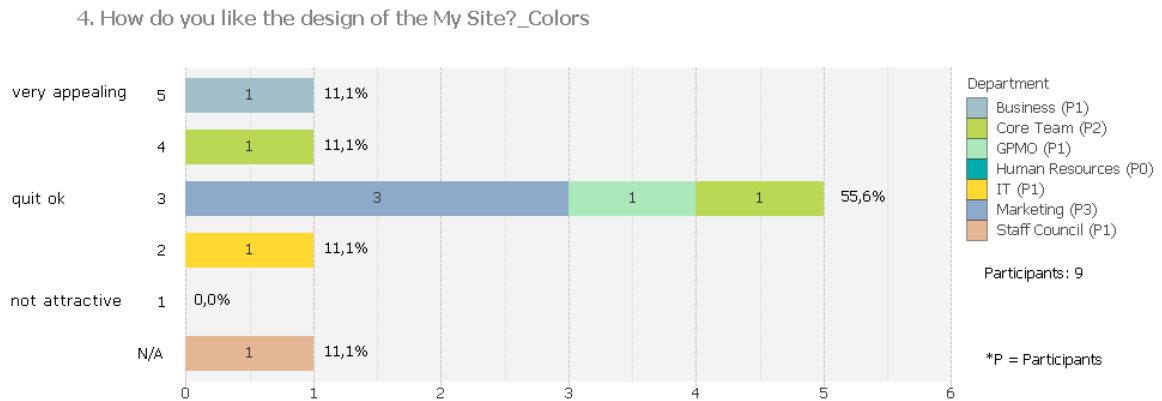
Figure 27: Dashboard - Rating

Figure 28 shows a more detailed view on the text area with an elaborated description of the changed area below:

2. Answer text: A rating in SharePoint is based on numbers starting from one to x, where x is a freely chosen number. This numbers are ranked according to their meaning. Therefore this meaning needs to be presented in the chart. SharePoint saves the meaning of the range using three words. One for the lowest, one for the mean and one for the highest value in order to understand the meaning of the numbers. These words are displayed directly next to the answer numbers.
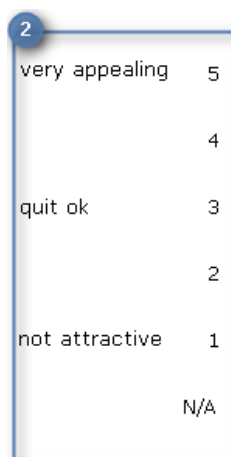
Figure 28: Dashboard – Rating visual areas

7 Summary & Results

In the previous chapters two different approaches for viewing and analyzing surveys were discussed, which were created and processed using SharePoint 2010. The first approach was to analyze the data directly with SharePoint. The second one was to analyze the data using an external application which receives the data via Spreadsheet export from SharePoint. This approach was done with QlikView due to reasons I already mentioned in Chapter 5.1.

This concluding chapter should describe and elaborate the benefits and also the drawbacks of both solutions. Eventually both results will be compared with each other to provide a sound statement about the suitability of each approach in reference to the purpose.

7.1 Survey analysis with SharePoint

As described in Chapter 4.2 SharePoint offers three different analysis options, which are the following: (1) all responses, (2) a graphical summery and (3) exporting the data to a spreadsheet. The advantages and disadvantages of these different approaches will be discussed in more detail below.

7.1.1 All Responses

All Responses is a simple list of answers group according to the response of each person .More details on this view are to be found in chapter 4.2.1.

Pro:
- Presents all data evaluated by the survey and participant.
- The presentation fits every type of data set.
- Good for detailed information on one person.

Con:

- No graphical Report.
- No relationship between other responses.
- No general statements.
- No general view on all survey data at once.
- No anonymity. (the participants names are visible)

The (1) all Responses approach is suitable if detailed information on a particular participant is required.

7.1.2 Graphical Summary

The second view is the most powerful view in SharePoint because it compares all answers related to the question, calculates several basic numbers and generates a graphic. Further details are listed in chapter 4.2.2.

Pro:

- Clear outline on where the most and less answers were provided
- Anonymous
- Presents all answers corresponding to a question.

Con:

- No link between the user and a user group like a department. There is no way to make a statement on how a specific user group has responded to that survey.
- No individual color highlighting.
- Many different answers lead to crowded and confusing graphical response.
- Adds no additional value to the SSQC Plain text, Numbers, Single Lookup and Multiple Lookup if the answer set is too broad.
- Offers only information on question level, no general survey analysis is possible.

Depending on the type of class the data belongs to, this graphic provides more or less meaningful results. If a fixed set of predefined answers is available, these graphical results can add additional value to the analysis. Which means further, with a dataset where almost every participant answered individually, the calculated numbers will have no additional meaning.

The example in Figure 29 illustrates that the graphical summary can add no value to question analysis when it comes to handwritten, individual answers except that it displays all answers corresponding to the question.

Figure 29: mySite – Getting Started – Question 8a (Graphical Summery)

7.1.3 Export to Spreadsheet

This export to spreadsheet option exports the data collected by the user and stores it in format of an Excel file. The ensuing analysis of the data is not done in SharePoint but, for example, in Excel. The big drawback of this approach is that only the collected answers are exported, which means further no additional Meta data is available through this function.

Pro:
- Export the answers out of SharePoint to analyze it with another tool.

Con:
- Not anonymous (the participants names are visible)
- Only the data entered by the user is presented.

- No Meta data information like for example question type, sort order, set of answers, etc. is included in this export.

7.1.4 Summary

SharePoint offers enough information to get a good first impression of the survey data. Especially the graphical view is very useful for single choice or multiple choice questions. The spreadsheet export offers an overall view on the data and the all responses view opens a detailed and formatted view on selected users.

On the other hand no general overview for the whole survey is available. Furthermore is also not possible to group the participating people into something with additional value like departments. Therefore a deeper analysis of the data requires a more detailed look into almost every answer of every user. This is why there is no way to preserve the anonymity of the user. The analyst knows exactly who responded how on what question. Moreover, the fact that the analyst needs three separate analyses to do a whole evaluation can be very time consuming.

Subsequently it is possible to analyze the survey only using features SharePoint is providing, but only on condition that a small dataset has to be analyzed. For a number exceeding 20 participants it can be very frustrating and time consuming to rummage through all the data.

7.2 Survey analysis with QlikView

The QlikView SPSurvey Dashboard described in chapter 6.2.4.4., provides all the functionalities presented by the SharePoint views and the spreadsheet output as well, but also includes some new functions:

- Answers per department (Chart and List)
- Participants
- Filter and selection boxes for questions, answers, surveys and SSQC's

Apart from the filters and charts presented on one single dashboard, QlikView adds with its in memory computing technology a fast and easy solution to analyze the data in multiple ways. The QlikView application combines all the above functionalities in one dashboard. Additionally the analysis can be extended on the frame of a question and even multiple surveys can be analyzed in a combined context.

On the downside analyzing the data with QlikView means to use the spreadsheet output from SharePoint. This output, as it is described in Chapter 4.2.3, lacks in important Meta information needed to program a dashboard like it is defined in Chapter 6.2.4.4. In order to do that, comprehensive information must be collected through the Survey Creator. This additional information has to be stored and saved be the Survey Creator himself. The best way to fulfill this task is to copy the data into the file surveysettings.xlsx already during the creation of a survey. But the Survey Creator only has to modify the Excel file, no QlikView scripting or changes on the dashboard is necessary.

Pro:

- Analysis of the data not by user but by department
- Anonymous
- Covers the same functionality as the SharePoint surveys and adds even more functions to it
- Fast and easy to use

- Possibility to analyze data from various perspectives by using different filters
- A Color space, consistent throughout the application
- All information available within one dashboard
- Only the file surveysettings.xlsx needs modifications - no particular QlikView knowledge is necessary to load the application.
- Analyses of surveys in a combined context

Con:
- Survey creator has to enter data manually into the surveysettings.xlsx to make the data load work
- More preparation effort is required
- Analyzing the application is not possible directly after the data export from SharePoint due to the additional task of creating the surveysettings.xlsx file

7.3 Conclusion

The BI (Business Intelligence) tool QlikView offers the analyst one consolidated frontend regardless which SSQC type the data is from. Furthermore all functionality offered by the SharePoint views is covered within this analysis with additional new functions. Therefore it seems by far the more suitable tool to analyze SharePoint surveys. The negative aspect of the SPSurvey QlikView application is the need of additional information, manually stored out of SharePoint in order to make the data load of the application work.

In short, it really depends on the complexity of the application, the preparation time and the timeframe available to decide which tool is suitable for analyzing the data.

An approach could be that SharePoint analysis always offers a good first start to get a felling on how the data might look like independently of the general conditions of the surveys. Going into more detail, SharePoint is still a suitable solution for less complex surveys with participants less than 20 people. Every other purpose, especially when

multiple surveys must be analyzed, the extra effort is well spent in order to make broader analysis options in QlikView available.

8 Outlook

SharePoint 2010 is not the latest version of the Microsoft SharePoint line-up. The latest SharePoint version was released on the 11 October 2012 and is called SharePoint 2013. (McNulty 2012) SharePoint 2013 has changed in many ways compared to 2010. (Microsoft, Funktionen und Features) However, taking a closer look on what has changed regarding surveys in 2013 only one new feature stands out. The Excel survey now enables users to generate surveys out of office 365 with Excel. Unfortunately it doesn't extend the already existing survey features of SharePoint and neither does it widen the analysis possibilities. (Microsoft, Surveys in Excel) In fact regarding administrating and analyzing nothing has changed at all, compared to 2010. (Microsoft, Create a survey)

That means that the survey analyzing features have not been and will not be enhanced by Microsoft in the near future, because so far there is no evidence that Microsoft intends to pay more attention to this part. Therefore the QlikView SPSurvey Dashboard is still compatible with SharePoint 2013 and probable further releases in the near future.

8.1 Enhancement of the QlikView application

The current release of the QlikView SPSurvey application is not at an endpoint and is still extendible in its capabilities. Generally the application can be further enhanced and modified in two main parts: (1) concerning the functional side and (2) concerning the application load and how the survey data has to be maintained.

The following subchapters describe how the application can be enhanced in more detail.

8.1.1 Enrichment of the application

From a functional perspective the next step would be to enhance the application in order to enable loading all SSQC into the QlikView application. That means not only to provide reload but also to enhance the applications dashboard in a way so that the attributes of the new data can be used as enhancement to the value of the discovery.

8.1.2 Simplification of the data load

Currently, it is necessary to enter a lot of information from SharePoint into the file surveysettings.xlsx manually by the Survey Creator. The goal would be to reduce the amount of input to zero or almost null. This could be accomplished by enhancing the application in two ways:

- Automated department assignment
- Import survey Meta data directly from SharePoint

8.1.2.1 Automated department assignment

One of the most important information is the allocation of the user to a department. All users participating in this surveys need to be classified referring to the department they belong to.

Most companies already have a lot of information on users in their directory service and QlikView for example is able to read the Active Directory from Microsoft. (Blang 2012) So the handmade assignment could be automated by using information directly from the directory service.

8.1.2.2 Survey Meta data import directly out of SharePoint

SharePoint 2010 provides three API's (application programming interface) to access data apart from actually using the SharePoint frontend: (Microsoft, SharePoint 2010 API)

- Server-side Object Model

- Client Object Model
- Web Services (SharePoint 2010 Web Services or the REST SharePoint interface (Venkat 2011))

This enables subtraction of data directly from SharePoint and storing it elsewhere. The best solution would be to access SharePoint directly with QlikView using a web service or REST interface because QlikView is able to access web services. (Stoichev 2010) If this approach is not working with SharePoint 2010 it is possible to use some kind of ETL tool or application to withdraw the data first and then store it in a readable format for QlikView to load it afterwards.

9 Bibliography

AIIM. *About AIIM.* n.d. http://www.aiim.org/About (accessed December 10, 2013).

—. *What is Enterprise Content Management (ECM)?* 2013. http://www.aiim.org/What-is-ECM-Enterprise-Content-Management (accessed December 02, 2013).

Awad, Elias M., and Hassan M. Ghaziri. *Knowledge Management.* Delhi, India: Dorling Kindersley (India), 2007.

Blang, Rutherford Le. *Connecting to and Querying Active Directory for Users.* April 03, 2012. http://community.qlikview.com/thread/50661 (accessed December 4, 2013).

Chan, Ivy, and Chee-Kwong Chao. "Knowledge Management in Small and Medium-sized Enterprises." *Commun. ACM*, April 2008: 83-88.

Eggert, Sandy. *Enterprise Content Management.* Berlin: GITO mbH, 2007.

Fenz, Stefan. "Increasing knowledge capturing efficiency by enterprise portals." *VINE*, 2012, 2 ed.

Gilbert, R. Mark, Karen M. Shegda, Kenneth Chin, Gavin Tay, and Hanns Koehler-Kruene. "Gartner: Magic Quadrant for Enterprise Content Management." October 18, 2012. http://www.project-consult.de/files/Gartner%20Magic%20Quadrant%20for%20Enterprise%20Content%20Management%202013.pdf (accessed November 28, 2013).

Griffith, T.L., J.E. Sawyer, and M.A Neale. *Virtualness and knowledge in teams: Managing the love triangle of organizations, individuals, and information technology.* MIS Quarterly, 2003.

Janssen, Cory. *In-Memory Computing.* March 12, 2013. http://www.techopedia.com/definition/28539/in-memory-computing (accessed December 03, 2013).

Lambert, Bob. *A QlikView QuickStart: first steps for learning QlikView desktop.* 02 20, 2012. http://robertlambert.net/2012/01/a-qlikview-quickstart-first-steps-for-learning-qlikview-desktop/ (accessed 11 06, 2013).

Manhart, Dr. Klaus. *Business Intelligence (Teil 3): Datenmodellierung – Relationale und Multidimensionale Modelle.* 02 05, 2008. http://www.tecchannel.de/server/sql/1744994/business_intelligence_teil_3_datenmo

dellierung_relationale_und_multidimensionale_modelle/index11.html (accessed 11 06, 2013).

Mayer, Host Otta. *Interview und schriftliche Befragung: Grundlagen und Methoden empirische Sozialforschung.* München: Oldenbourg Verlag, 2013.

McNulty, C. *SharePoint 2013 Releases to Manufacturing (RTM) 10/11/12 @cmcnulty2000 @qsharepoint #sp2013 - KnowPoint.* October 12, 2012. http://www.chrismcnulty.net/blog/Lists/Posts/Post.aspx?ID=59 (accessed December 4, 2013).

Microsoft. *Create a survey.* n.d. http://office.microsoft.com/en-us/sharepoint-help/create-a-survey-HA102775328.aspx (accessed December 04, 2013).

—. *Deciding Which SharePoint 2010 API to Use.* n.d. http://msdn.microsoft.com/en-us/library/hh313619.ASPX (accessed December 4, 2013).

—. *External Data in SharePoint 2010.* n.d. http://msdn.microsoft.com/en-us/library/ff798503.aspx (accessed December 04, 2013).

—. *Funktionen und Features in SharePoint 2013.* n.d. http://technet.microsoft.com/de-de/sharepoint/fp142374.aspx (accessed December 4, 2013).

—. *Introduction to managed metadata in SharePoint Server 2010.* n.d. http://office.microsoft.com/en-us/sharepoint-server-help/introduction-to-managed-metadata-in-sharepoint-server-2010-HA101859256.aspx (accessed December 04, 2013).

—. *Surveys in Excel, hosted online.* n.d. http://office.microsoft.com/en-us/web-apps-help/surveys-in-excel-hosted-online-HA103171060.aspx?CTT=1 (accessed December 4, 2013).

Miles, Dog. "State of the ECM Industry 2011: How well is It meeting business needs?" *http://www.aiim.org/.* March 21, 2011. http://www.aiim.org/pdfdocuments/IW_ECM_State-of-Industry_2011.pdf (accessed November 28, 2013).

Morgen, Jacob. *The Collaborative Organization: A Strategic Guide to Solving Your Internal Business Challenges Using Emerging Social and Collaborative Tools.* United States of America: McGraw-Hill, 2012.

OPTIMAL SYSTEMS GmbH. *Definition von Enterprise Content Management.* n.d. http://www.optimal-systems.de/ecm-einstieg/was-ist-ecm.html (accessed December 19, 2013).

O'Reilly, Tim. *What Is Web 2.0.* O'Reilly Media, Inc, 2009.

Pandre, Andrei. *Comparison | Data Visualization.* 2010. http://apandre.wordpress.com/tools/comparison/ (accessed December 03, 2013).

—. *Updated Comparison of Data Visualization tools | Data Visualization.* December 18, 2011. http://apandre.wordpress.com/2011/12/18/dv-comparison-2011/ (accessed December 03, 2013).

Passionned Group. *What is Business Intelligence (BI)?* n.d. http://www.passionned.com/business-intelligence/what-is-business-intelligence-bi/ (accessed December 19, 2013).

Richardson, Adam. *Inventing the Collaborative Workspace.* November 21, 2011. http://blogs.hbr.org/2011/11/inventing-the-collaborative-workspace/ (accessed December 4, 2013).

Richter, Wolfgang. *Star-Schema.* 10 1, 2001. http://www.iicm.tugraz.at/cguetl/education/thesis/wrichter/thesis-final/node44.html (accessed 11 6, 20113).

Rouse, Margaret, and Stan Gibilisco. *scatter plot.* December 2012. http://whatis.techtarget.com/definition/scatter-plot (accessed December 04, 2013).

Schaffner, J., B. Eckart, D. Jacobs, C. Schwarz, H. Plattner, and A. Zeier. "Data Engineering (ICDE), 2011 IEEE 27th International Conference on." *Predicting in-memory database performance for automating cluster management tasks.* http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5767936&isnumber=5767827, April 11-16, 2011.

Stoichev, Stefan. *Using QlikView with Web Services.* 09 20, 2010. http://blog.qvapps.com/2010/09/20/using-qlikview-with-web-services/ (accessed 11 13, 2013).

Venkat, M. *Introduction to the REST interface in Sharepoint 2010.* July 7, 2011. http://pravahaminfo.blogspot.co.at/2011/07/introduction-to-rest-interface-in.html (accessed December 4, 2013).

Weintraub, Alan, Craig Le Clair, and Cheryl McKinnon. "The Forrester Wave: Enterprise Content Management, Q3 2013." September 19, 2013. http://www.project-consult.de/files/Forrester%20Wave%20ECM%202013.pdf (accessed November 28, 2013).

Yurgosky, Patrick. *Beyond Dashboards: Business Intelligence Tools for Program Analysis and Reporting.* June 2012. http://www.idealware.org/articles/beyond-dashboards-business-intelligence-tools-program-analysis-and-reporting (accessed December 6, 2013).

10 Table of Figures

11 Table of Code Samples

12 List of Tables

13 List of Abbreviations

AIIM
    Association for Information and Image Management 5
API
    application programming interface 73
BI
    Business Intelligence 10
CMS
    Content Management Systems 4
ECM
    Enterprise Content Management Systems 4
etc
    and so on and so forth...54
KM
    Knowledge Management.. 7
N/A
    no answer 23
RGB
    red, green and blue 38
SSAS
    SQL Server Analysis Services..28
SSQC
    SpreadSheet Question Classes 19
WEB
    World Wide Web 4
YPP
    Young Professional Potentials..13