CHALMERS

Smarter Software Release Planning

A Case Study on Using Business Intelligence and Social Media Mining to Analyze Market Trends

Master of Science Thesis in Software Engineering and Technology

EVELINA DAVIDSSON
SARAH GEAGEA

Department of Computer Science and Engineering
Chalmers University of Technology
Göteborg, Sweden, June 2012

The Authors grants to Chalmers University of Technology the non-exclusive right to publish the Work electronically and in a non-commercial purpose make it accessible on the Internet. The Authors warrants that he/she is the author to the Work, and warrants that the Work does not contain text, pictures or other material that violates copyright law.

The Authors shall, when transferring the rights of the Work to a third party (for example a publisher or a company), acknowledge the third party about this agreement. If the Authors has signed a copyright agreement with a third party regarding the Work, the Authors warrants hereby that he/she has obtained any necessary permission from this third party to let Chalmers University of Technology store the Work electronically and make it accessible on the Internet.

Smarter Software Release Planning - A Case Study on Using Business Intelligence and Social Media Mining to Analyze Market Trends

EVELINA DAVIDSSON
SARAH GEAGEA

© EVELINA DAVIDSSON, June 2012.
© SARAH GEAGEA, June 2012.

Examiner: LARS PARETO

Chalmers University of Technology
Department of Computer Science and Engineering
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Department of Computer Science and Engineering
Göteborg, Sweden, June 2012

Smarter Software Release Planning – A Case Study on Using Business Intelligence and Social Media Mining to Analyze Market Trends

Evelina Davidsson, Sarah Geagea
Chalmers University of Technology
Department of Computer Science and Engineering
{evelinad, geagea}@student.chalmers.se

Abstract

BACKGROUND: *Tracking market trends is an important part of market-driven software release planning, although often challenging because of rapid changes in consumer demand. Recently, there has been a growing interest to utilize social media mining to track and analyze consumer opinions. A Business Intelligence tool for analyzing buzz on social media could potentially give valuable decision support for software release planning by providing up-to-date insights into market trends.*

METHOD: *This paper presents a Business Intelligence tool which was developed using a design research approach. The application serves as a proof of concept for analyzing buzz on social media and was evaluated in the context of software release planning and market-driven planning separately through a case study.*

RESULTS: *During evaluation it was found that buzz on social media was considered to be an interesting source of information, but that specific applications within software release planning were not obvious. An information quality assessment confirmed big challenges in interpretation of data from social media, which could partly be addressed with increased domain knowledge and search optimization.*

CONCLUSIONS: *The case study confirms that a Business Intelligence tool for mining social media could serve as a complementary support for decisions regarding software release planning. Recommended future work includes testing the hypothesis in a pure market-driven software release planning context, as well as incorporating geographical attributes.*

KEYWORDS: Software Release Planning, Social Media Mining, Business Intelligence, Information Quality Assessment

1 Introduction

As the rate of technology change is increasing, product planning within software development is becoming more important, and challenging, in order to gain and sustain a competitive edge. In market-driven software development, where there is no commissioning customer, keeping up-to-date with consumer demands is especially crucial. Additionally, end-customers are getting more powerful because of global competition and information transparency [15], which places great pressure on quality as well as time to market.

Software release planning is the process of selecting sets of features for sequential product releases in such way that product value is optimized with respect to given constraints and resources [48]. It is known as a complex decision-making activity because of the many factors involved [48, 24, 35] and because of the challenge of maintaining an appropriate balance between technology push and market pull [41]. The act of determining long-term visions for a product, often referred to as strategic release planning or roadmapping, must also be supported by release planning on a more detailed, operational level [39, 50]. There are several systematic methods for release planning, as summarized by Svahnberg et al. [54], and the need for even more sophisticated decision support for this process has been recognized [21]. However, ad hoc approaches are still common in industry.

Although soft factors, i.e. product and customer value, are key drivers for release planning in market-driven development, there seems to be more focus on hard factors such as technology, cost, time and dependency constraints in existing models for release planning [54]. In some cases, workshops are used in release planning to brainstorm about customer demands [41, 31], but subjective decision-making and groupthink are always potential risks in isolated groups [22]. According to SWEBOK[1], software engineers often need to rely on a marketing unit to provide objective information about consumer

[1] Software Engineering Body of Knowledge

demands [1]. However, traditional market research and customer surveys take time to conduct and only target a subgroup of potential customers. It might also be the case, especially within big companies, that reports provided by the marketing department contain information on an organizational level, but is not adjusted to the needs of release planning for specific products. In addition, Karlsson et al. [28] found that there is often a gap between the marketing department and software engineers.

One interesting and growing source of consumer opinions is social media. During the last years, the use of various social media platforms has drastically increased [4]. Facebook, which is one of the most popular platforms and was founded in 2004, reached 845 million active users globally in 2011 [14]. The microblog[2] Twitter reached 140 million active users in March 2012 and had an average of 340 million tweets posted per day at the time [58]. These types of platforms enable people to easily share and spread their opinions about all kinds of topics.

A study made by AYTM[3] [3] in 2011 shows that social media users are, among other things, expressing their opinions about brands, which might be of interest for software release planning purposes. Another study, comparing different sources for product information, found that people tend to trust friends and online sources to a greater extent than traditional channels such as TV commercials and news sources [39]. This suggests that analyzing social media content can contribute to valuable insights about consumer demands in a market-driven context, as a complement to traditional market research.

However, the large amount of user-generated content available on social media platforms makes it difficult to get an overview of trends. The activity of structuring and visualizing big sets of data to facilitate decision making is often referred to as Business Intelligence (BI) [43]. With roots back in the 60s, BI is a mature field and is widely used in industry [43]. While traditional BI solutions are built upon company-internal data (e.g. sales history, inventory and purchases), external sources are becoming increasingly common [51]. Especially, social media mining is an emerging and popular area often connected to BI.

Because of rapid changes in both technology and market, another key challenge in release planning is keeping the plan alive and up-to-date [41, 24]. A BI solution for the purpose of continuously tracking market trends based on social media content could therefore be a good complement to traditional market research. Allowing a product planner to directly access this type of information, as opposed to going through a marketing department, can offer the flexibility and level of detail needed within release planning.

Our hypothesis is that a BI application, as described above, could potentially contribute to smarter release planning in market-driven software development by providing up-to-date and objective information about consumer demands as input to more informed discussions when prioritizing requirements.

The remaining part of this report describes an attempt to investigate and evaluate the stated hypothesis and is structured as follows: chapter 2 specifies delimitations for the project, whereas chapter 3 and 4 provide sufficient background to the different subjects involved as well as related work. Further, the research design is explained in chapter 5, followed by obtained results in chapter 6 and main conclusions in chapter 7.

2 Delimitations

In an ideal case, the hypothesis of this study would be evaluated in a pure market-driven software release planning context. However, such case could not be obtained for this study, given that the focus lied on developing a working prototype within the limited time frame of this thesis. Instead, this study includes an evaluation case that covers the perspective of a business-to-business software product planner and an auxiliary case, which evaluates the hypothesis in a market-driven context.

3 Background

This study incorporates a wide range of subjects from different fields, hence the reader might not be familiar with them all. The following sections present a short background to each area in order to give sufficient context for the study.

3.1 Software Release Planning

In Software Engineering, release planning comprises the assignment of requirements to specific product releases. It is often described in literature as the process of selecting optimal sets of features to include in each release, with regards to given constraints [48, 7]. More specifically, release planning is concerned with finding the appropriate balance between stakeholder priorities and development resources such as available technology and competence [48, 7]. It can be noted that one of the main stakeholders in market-driven software development is the group of potential customers within a targeted market.

Release planning can be performed on a high-level to create a long-term vision and is in this context often called strategic release planning or roadmapping [41, 48, 2]. The different approaches

[2] A microblog is a blog with short posts. Specifically on Twitter, the messages are limited to 140 characters [25].
[3] Ask Your Target Market

on this level can generally be described with the terms technology push, i.e. exploration of technology and its opportunities, and market pull, i.e. defining products based on existing customer needs [41]. The result of roadmapping is often a graphical presentation of how resources will be used in product offerings to meet market demands with respect to a given time plan [41].

The long-term vision created in strategic release planning must also be supported by release planning on a more detailed, operational level [48, 2]. This type of short-term release planning includes prioritizing specific requirements into upcoming releases and allocating tasks appropriately [2, 10]. Operational release planning is sometimes referred to as requirements prioritization or triage [10].

Release planning is widely recognized as a complex or wicked problem because of the many challenges associated with the process [24, 48, 7]. When defining the scope for each release, trade-offs between time, cost and quality attributes must be made [10, 38]. This can be difficult because of competing stakeholder opinions [38], requirement interdependencies [48, 7, 10] and continuous changes in environment such as market demands [24, 48]. Additionally, insufficient information regarding effort and resources can result in big uncertainty within release planning [24, 48].

As summarized by Svahnberg et al. [54], existing literature suggests several methods to support software release planning. One example of a systematic model is the Cost-Value approach, which is used to compute the best requirement candidates based on interpreted value and estimated effort [27]. EVOLVE is another approach combining computational strength with human common sense and hence can be categorized as a hybrid method. Non-systematic, ad hoc approaches to release planning are, however, still common in industry, even when developing large products [49].

3.2 Social Media

Social Media can broadly be described as online platforms that include user-generated content [62, 45, 30]. These platforms come in many forms (e.g. blogs, social networking sites, content communities, virtual gaming worlds etc.) [62] and employ web-based technologies that enable communication and interaction amongst individuals through text, video and audio [45, 30].

Social media differs from traditional media, such as newspapers and books, in the sense that almost anybody can publish or access non-reviewed information [62]. Moreover, social media and social networks are often thought of as being the same thing, although social networks are only a subset of social media [62]. More specifically, a social network is defined as a social structure that typically consists of nodes, such as individuals, and links representing the relationship between nodes [62].

Kietzmann et al. [30] present a framework that aims to describe the building blocks of social media. It is intended to be used both for developing social media platforms and for guiding firms when engaging in social media activities. The framework is illustrated as a honeycomb and includes seven different blocks, i.e. social media functionalities. The first one is *identity*, which is concerned with the degree to which users share information about themselves, whereas the second block, *presence*, refers to the degree to which users know that other users are available. Further, the degree to which users are associated to one another appertains to the *relationships* block, while the degree to which users know of other users' trustworthiness is described under the *reputation* block. The remaining blocks include *groups*, *conversations* and *sharing*, which refer to the degree to which users form communities, the degree to which users communicate with each other and the degree to which users share and acquire content [30].

Today's major social media sites fall under the category of at least three honeycomb blocks, even though some of those blocks might be more evident than others [30]. Facebook, for instance, is primarily concerned with relationships. This is mainly because the platform allows users to determine which other users to be friends with, i.e. have a connection to, and categorize them according to the relationship they have in real life, e.g. relatives, classmates, co-workers etc. Further, users also have the possibility to decide with whom to share their information (e.g. status updates, pictures, links and videos). Twitter, on the other hand, is more focused on conversations since its main functionality comprises sharing 140 character long messages, also known as Tweets, which are publically available to anybody in the network.

3.3 Business Intelligence and Data Mining

The term Business Intelligence refers to technologies, tools and methods to extract valuable information from large volumes of data in order to enable better decision making. This includes practices for collecting, integrating and analyzing data as well as finally visualizing relevant information [11]. BI has roots back in the 60s, when companies began using simple reporting on transactional data, such as "sales by region" [51, 43]. Transactional data of this kind is commonly stored in relational databases within ERP[4] systems and is often referred to as structured data.

Over time, the need for more complex reporting emerged as well as analysis of more unstructured

[4] Enterprise Resource Planning

data. Today, an estimate of only 20% of an enterprise's information assets is structured while 80% is unstructured, mainly in the form of text documents such as contracts, policies and knowledge management reports [40, 11]. Traditionally BI has also been about analyzing company-internal data, but external data such as market surveys and web sources are increasingly being incorporated into companies' BI solutions [51].

Often connected to BI is the field of data mining, i.e. automated discovery of patterns and knowledge from large data sets. The different data mining techniques can generally be divided into two categories: descriptive data mining, i.e. characterization of data, and predictive data mining, i.e. performing induction on current data to make predictions [19]. For instance, unstructured data such as text documents can be characterized using text mining techniques [40].

The increase of online data sources, which most often includes unstructured data, have led to a sub field of data mining called web mining, i.e. the use of data mining techniques to extract information from web documents and web related data [32, 8]. Web mining can also be further divided into three categories based on the characteristics of data sources; mining of web content, web usage and web structure [32]. One example of web usage mining is the search statistics offered by Google Trends[5].

3.3.1 QlikView

QlikView[6], a self-contained BI product by QlikTech, is identified by Gartner as one of the leading BI platforms on the market along with players like IBM, Microsoft and Oracle [18]. Characteristics for QlikView include wholly in-memory data storage as well as an intuitive user interface [18]. Data can be collected via SQL like scripts from multiple sources, e.g. databases, Excel documents and web pages, and transformed into an associative model [42]. Once data is loaded into a QlikView application, it can be visualized through customized dashboards. The underlying associative model enables the user to do further analysis by interactive aggregation and filtration of the data.

3.4 Social Media Mining

With the increasing amount of user-generated content on social media platforms, an extension of web mining under the name social media mining has become a popular field, attracting researcher from several areas [36, 32]. Although social network analysis has a long history in the field of sociology [6, 50], the emerging possibilities of using web mining techniques on online social networks has opened up research opportunities for a variety of disciplines including internet computing, social computing, marketing, business, pattern recognition and artificial intelligence [32, 6]. Not least within BI, the potential of mining user-generated data has been recognized, among others by Gartner [16].

The different categories of web mining mentioned in section 3.3 can also be applied to social media mining [32]. Social media content mining is focusing on the user-generated content on social media platforms including text, images and videos. Social media activity and social media relations mining are, similarly to web usage and web structure mining, concentrated on the behavior of and relation between users respectively [32]. One example of an existing commercial product that is based on all three categories is the Facebook application Wisdom[7] which offers statistics on content, activity and relations of Facebook users.

Although the field of social media mining is still immature and no key metrics have been established [16], many have suggested metrics regarding how much content there is on a certain subject, attributes of this content and the characteristics of the content authors. The metrics in table 1 can be used to quantify this kind of information.

There are currently several open APIs available for fetching user-generated data from social media such as mentions on a specific subject. Some

Metric	Description	Example
Mentions	Number of occurrences of a specific subject.	*E.g. there is an average number of 1509 tweets/day containing the word "Android".*
Sentiment	A number representing the attitude of the content, often in the form of positive, negative or neutral.	*E.g. the average sentiment of the tweets in previous example is 0,1 on a scale from -1 (negative) to 1 (positive).*
Influence	A number representing the influence of a social media user, based on how many other users are reached and affected by his or her content.	*E.g. Bill Gates has an influence score of 76, on a scale from 0 to 100 (according to the influence measure Klout).*

Table 1 - Examples of social media metrics

[5] Google Trends: http://www.google.com/intl/en/trends/about.html#7
[6] QlikView: http://www.qlikview.com/
[7] Wisdom: http://www.wisdom.com/

examples are the Twitter Search API[8], the Facebook Graph API[9] and the Twingly Analytics API[10]. For characterization of the gathered content, text mining algorithms can be used to compute sentiment scores. There are several commercial sentiment algorithms available on the market, but also some non-commercial. One non-commercial example is the Twitter Sentiment algorithm[11], which essentially uses an approach where words in the analyzed tweet are compared with pre-defined, negative and positive keywords [17]. In every tweet, the number of positive and negative keyword occurrences are counted, where the higher count determines the sentiment of the text. Another approach for text analysis is presented in the widely cited paper by Turney [57], where sentiment is determined by comparing the extent to which relevant phrases from a text co-occur with pre-defined positive and negative keywords. This approach is based on the hypothesis that positive and negative words, respectively, often co-occur by chance. Although there currently exist several text analysis approaches, the task of interpreting social media (or other short) messages is considered to be a challenging task [9, 61].

Furthermore, influence of users on social media platforms can be computed based on reach and effect on other users, e.g. numbers of followers and retweets on Twitter. A relatively well known influence measure for users on social media platforms is Klout score[12].

3.4.1 QVSource

The software tool QVSource[13] offers several connectors for fetching, analyzing and integrating data from different social media platforms into QlikView applications. The connectors make use of social media platform APIs as well as text analysis APIs and Klout score via http requests.

3.5 Information Quality Assessment

The quality of a decision is highly dependent on the quality of the underlying information [52]. With today's large amount and diversity of available data, which might originate from uncontrolled sources, it is becoming harder to separate relevant data from noise or "garbage" [52].

The importance of assessing information quality (IQ) is widely acknowledged [52, 34] and has led to the development of several frameworks. One of them, presented by Wang and Strong [59], focuses on investigating the aspects of IQ that are important for information customers, i.e. the people in an organization who use the information. The framework divides IQ into the four categories presented in table 2.

Category	Description
Intrinsic	How accurate or believable the content is.
Contextual	The relevance and completeness of the content.
Representational	The interpretability and presentation of the content and how easy it is to understand.
Accessibility	The availability and accessibility of the content, provided by the information system in use.

Table 2 - Categories of Information Quality

The categorization above has served as a base for several other frameworks where one example is AIMQ [34], which is an approach for measuring, analyzing and interpreting IQ. Another example, presented by Stvilia et al. [52], focuses on describing typical IQ problems. Further, Katerattanakul and Siau [29] used the categories to develop a tool for measuring IQ of websites.

4 Related Work

While several commercial tools for analyzing social media have emerged[14] as an effect of the increased amount of user-generated content, there is still little understanding on their potential applications to business. Also, existing research in the field of social media mining, although popular and interdisciplinary, tend to focus on problems and methods rather than business applications. As stated by Bonchi et al. [6], there is a clear gap between the techniques developed by the research community and their applicability to actual and specific business processes.

Market research is perhaps the field with most suggested applications of social media mining. Reputation monitoring, to know what customers and potential customers are saying about a brand and its competitors, is suggested as an interesting business application [23, 6]. Real time customer response to a live event, such as a television show, is investigated by Marasanapalle et al. [36], while Surma and Furmanek [53] suggest the use of social media mining to improve market response. Monitoring customer experience has been explored by many

[8] Twitter Search API: https://dev.twitter.com/docs/api/1/get/search
[9] Facebook Graph API: http://developers.facebook.com/docs/reference/api/
[10] Twingly is a search engine for blogs. Twingly Search API: http://www.twingly.com/products/data
[11] Twitter Sentiment: https://sites.google.com/site/twittersentimenthelp/
[12] Klout score: http://klout.com/corp/kscore
[13] QVSource: http://www.qvsource.com/

[14] Two examples of commercial tools for social media mining are: Sysomos: http://www.sysomos.com/, Attensity: http://www.attensity.com/home/

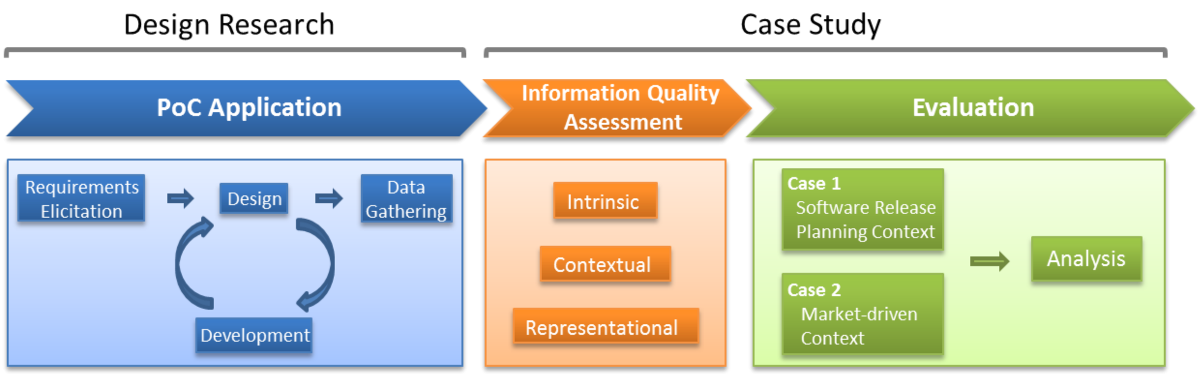


Figure 1 - Overview of research design

within the travel industry [20, 26, 63], a suitable context because users on social media willingly share their travel experiences [32].

Another area, which is considered to be one of the most important applications of social media mining techniques, is predictive modeling [6]. Possibilities for predicting, among other things, box-office revenues of movies [4, 62] and music album sales [12] have been identified. This due to findings that show positive correlation between how much buzz there is about an unreleased movie on Twitter, or music album on blogs, and how much revenue it brings in after it has been released. Strong correlations have also been found between the amount of tweets and election results [56, 62], making Twitter a valid source for indicating political opinion. Further, an example of a company recognizing the potential advantages of predicting the future by analyzing trends on the web, i.e. not only social media, is Recorded Future[15]. Currently, Recorded Future mines and analyzes data from over 40 000 different web sources in order to predict and provide insights into future happenings [44].

In Software Engineering, social media mining techniques for information gathering are not widely used. However, the need for decision making support in general within Software Engineering has been recognized, especially for requirements engineering and release planning [46, 47]. For instance, there exists methods such as Quantitative Win-Win and EVOLVE, as mentioned before, that focus on supporting decision-making regarding requirements negotiation and prioritization [46, 47]. Although both approaches value stakeholders' preferences, none of them consider the use of data or metrics derived from social media.

Other sources which have been mined for Software Engineering decision making purposes include patents [33], for identifying developing trends as input to roadmapping, and company-internal social networks, i.e. collaboration tools, in order to predict software build failures [60].

Regarding utilization of social media within Software Engineering, it is more commonly used for knowledge sharing and improving communication among software developers and teams [24, 5]. Further, social media tools for direct customer communication and feedback[16], have also emerged, making it easier to cooperate with customers. However, little has been found regarding the use of social media mining techniques in Software Engineering, especially for the purpose of supporting the release planning process.

5 Research Design

This report is partly based on a design research approach where a BI application, which serves as a proof of concept (PoC) for analyzing buzz on social media, was developed. Although similar applications exist today, it is hard to know the details and metrics lying behind them. Being involved in the creation of the application makes it easier to gain more control and gives a better understanding of the potential benefits and issues that might exist. Further, the remaining part of this report includes a case study where the PoC application was evaluated. First, the quality of the gathered social media data was assessed using a well-known information quality assessment framework. Second, the application was evaluated, using a qualitative approach, in two different companies with fairly separate contexts. To summarize, figure 1 illustrates the entire research design process used in this study.

5.1 PoC Application

The application was developed in cooperation with a global distributor of wireless products (from now on Company X). Although not working with software product development, Company X have similar challenges regarding market trend analysis as an input to purchase planning, i.e. deciding which products to invest in can be compared to deciding what kind of products to develop. In this sense, the decision making required from a purchase planner in

[15] Recorded Future: https://recordedfuture.com

[16] Some examples are
User Voice: http://uservoice.com/
Get Satisfaction: http://getsatisfaction.com/

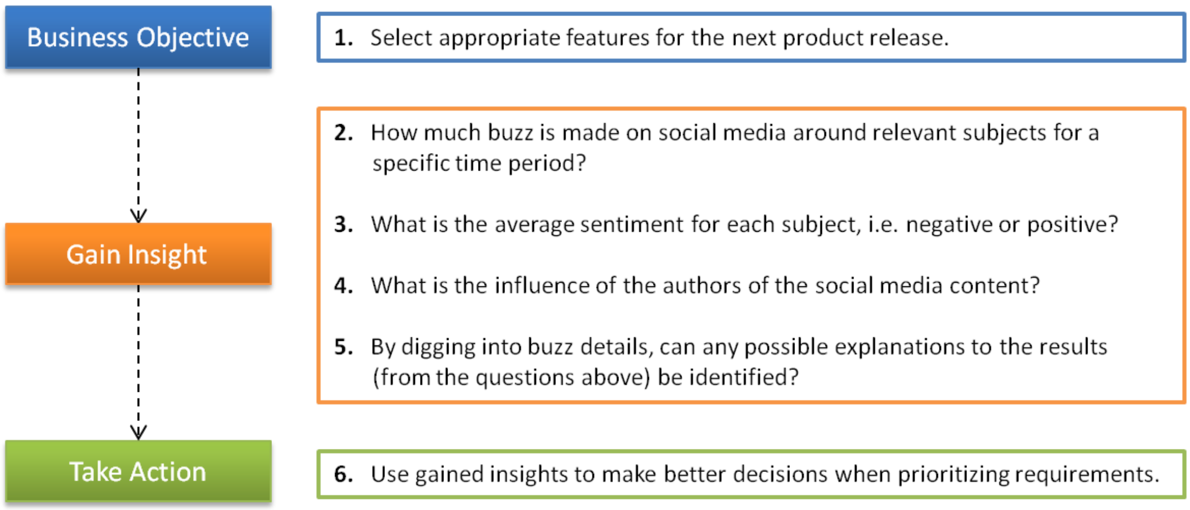


Figure 2 - Intended workflow for PoC application

a distributor company is similar to a product planner's decision making in a product development company. Furthermore, the company has a desire to explore social media as a source for BI and also possesses competence about BI in general and the BI platform QlikView in particular and therefore makes an appropriate candidate for our case study.

5.1.1 Requirements Elicitation

Initially, several interviews were held with possible stakeholders at Company X such as a Purchase Manager, Sales Manager and a Marketing Director (for a complete list of questions, see Appendix VIII). Most focus was put on the Purchase Manager since the decision making required from this role is similar as for a product planner as mentioned above. Additionally, a workshop was held with the BI team at Company X about possible usage of social media as a source for BI.

5.1.2 Application Workflow

As summarized by figure 2, the purpose of the PoC application is to support the process of software release planning by providing insight into consumer opinions, leading to more informed decisions about future product releases. A similar workflow could potentially be used in the context of purchase planning. Inspired by varied types of needs identified during requirements elicitation, the workflow was structured to support different levels of information aggregation. The basic idea was to first provide a quick summary of the information, i.e. mentions and average sentiment for each subject. Moreover, the application should provide the possibility to drill down to more detailed levels of information such as source and author characteristics. Ultimately, the user should be able to access the original texts for further analysis of interesting phenomena.

5.1.3 Data Selection

One important step in the development process was to decide upon which source or sources to gather data from. Since the purpose was to gain insights into market trends, social networks with a large user base where opinions can easily be shared were preferred. The blogosphere[17] was also considered since it serves as a rich data source, containing among other things consumer opinions and expert reviews. Hence, the sources presented in table 3 were chosen for the data sample.

Sources	Twitter
	Facebook
	Blogs (via the blog search engine Twingly)

Table 3 - Chosen social media sources for the study

Determining what subjects to search for from the sources listed above was another important aspect. Since content on social media is generated by humans it is crucial that the subjects that are searched for are human related and easily talked about in public in order to avoid biased results [62]. Also, the subjects for the PoC application were preferred to be relevant for the companies in which the application was evaluated. Since both of these companies are operating within the tele-communication industry, the subjects presented in table 4 were chosen.

Subjects	Search Terms
Mobile Network Operators	AT&T, T-Mobile, Vodafone
Mobile Operating Systems	iOS, Android, Windows Phone
Specific Upcoming Mobile Products	iPhone 5, Samsung Galaxy S3, Nokia Lumia 900

Table 4 - Chosen subjects for the study

The metrics chosen for this study are the ones described in section 3.4, i.e. mentions, sentiment and

[17] The blogosphere consists of all blogs and the connections between them.

influence. These simple and straightforward metrics were considered to give a good enough overview of the content on social media to serve as a starting point for investigating the hypothesis of this study. In order to present the chosen metrics, the data types listed in table 5 were fetched.

Data Types	Public Facebook posts and likes
	Tweets and retweets
	Klout score for Twitter users
	Blog posts
	Blog authority
	Sentiment

Table 5 - Data types fetched for the study

Further, two important limitations were introduced in order to manage the great amount of data as well as long processing times encountered for sentiment analysis. One of the limitations was to fetch posts and tweets written only by English users. This restriction was also due to the language limitations of the sentiment algorithm used for the application. The other limitation implied that data was only gathered for one hour per day. For this purpose, a small investigation was made to determine which hour of the day users on Facebook and Twitter are most active. An online article, published by Mashable [37], describes a study made by Virtue that investigates when Facebook users are most active. The study shows that the three largest peaks occur on weekdays at 11 a.m., 3 p.m. and 8 p.m. EST. Another study made by Sysomos [55] shows that the highest Twitter activity occurs between 11 a.m. and 3 p.m. EST. In conclusion, for the purpose of the PoC application, data was gathered for one hour per day around 3 p.m. EST.

Moreover, after the development of the application, data was collected during a two week period.

5.1.4 Technical Solution

The PoC application was developed in iterations, using the QlikView platform together with QVSource connectors. The connectors were used to fetch data from Twitter, Facebook and Twingly as well as sentiment and Klout scores. Details on the data model of the PoC application can be found in Appendix I and details on specific connectors used in the PoC can be found in the system architecture in Appendix II.

The system architecture also shows which source APIs the specific QVSource connectors use to fetch data. These APIs naturally set the limits for what is possible to do with QVSource. For instance, the Facebook, Twitter and Twingly APIs provide the possibility to fetch public data containing a specific search term. Although Twingly and Twitter are fairly easy to scan since the data to a great extent is public, their APIs are still limited in the sense that there is a restriction on available historical data. For instance, Twitter does not provide data older than one week and Twingly only returns a maximum of 1000 blog posts. Facebook is slightly different from the other sources since the data that is shared in the network can sometimes be private depending on how users have configured their privacy settings. However, Facebook provides the possibility to search through all public posts available without any time limit. Moreover, the Klout API is also limited since it only provides Klout scores for Twitter users. Hence, no influence measure was used for Facebook users and for blogs a similar authority measure, which is provided by Twingly, was used instead.

Further, the chosen sentiment algorithm for the PoC was Repustate[18] since this was the only algorithm provided by QVSource which used different analyzing techniques depending on the type of text, i.e. short or long text. This was considered to be a benefit since the data fetched from the different sources could be of both text types, e.g. long blog posts and short tweets.

5.1.5 Graphical Interface

In this study, the user interface of the application is not a primary focus since the aim is mainly to investigate whether information from social media can be useful for software release planning. However, some effort was placed into making an understandable user interface in order to facilitate the evaluation process. Firstly, a prototype of the user interface was made and evaluated together with the BI team at Company X. Based on the prototype and the different aggregation levels described in section 5.1.2, the application has been divided into five main tabs.

The first tab, which is also the first view when opening the application, presents a simple dashboard which aims to give an overview of the most important information, see chart in figure 3 (for the complete dashboard screen, see Appendix III). It includes three bar charts, where each chart represents one of the following subjects; operators, products or operating systems (e.g. figure 3 shows the bar chart for operators). Each chart presents the following:

- How much is being talked about each subject, i.e. mentions. This is indicated by the height of the bars.
- Whether the buzz is negative or positive, i.e. sentiment. This is indicated by the color of the bars, where the green/red color together with color intensity shows how positive or negative the buzz is.

The unit of the bar chart's x-axis represents time in the form of days. As seen in figure 3, there is also

[18] Repustate: www.repustate.com

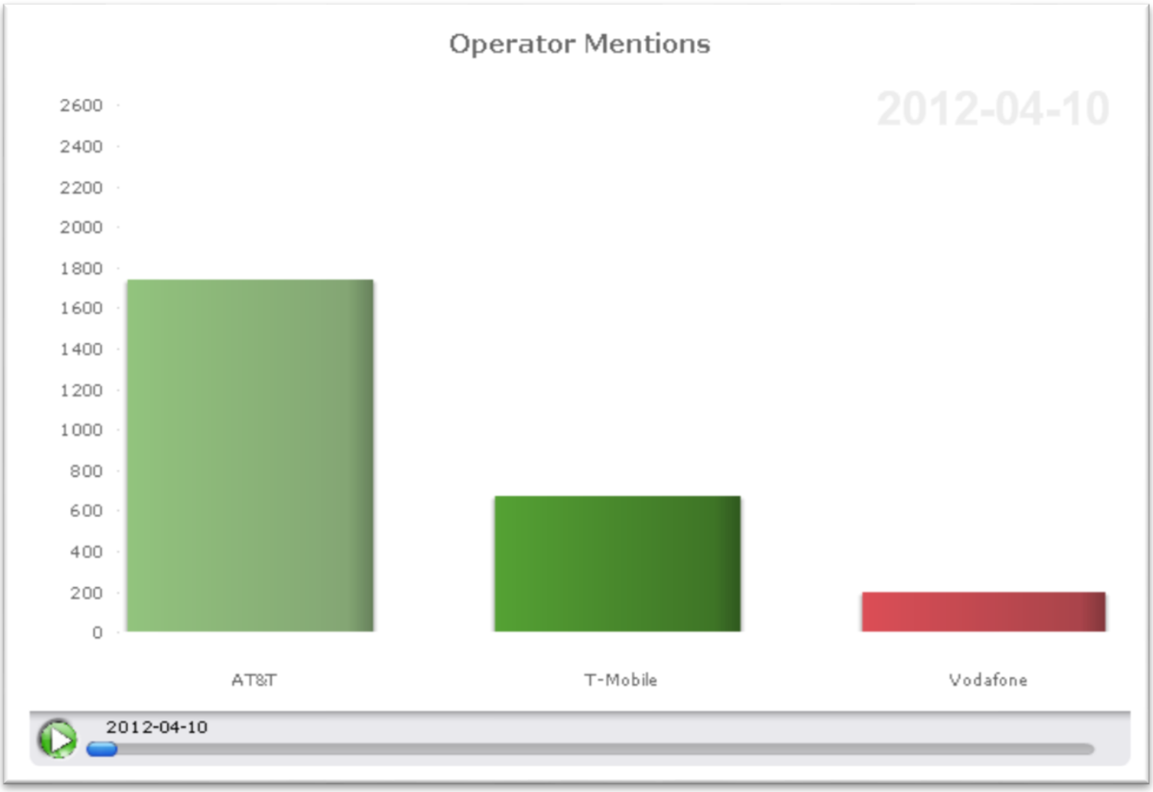


Figure 3 - PoC application screen shot: Dashboard chart

an option to play an animation of the bar chart to show how the buzz changes from day to day during a specific period of time. The calendar on top of the tab sheet allows a user to specify the time period of interest and is therefore included in every tab throughout the application. Further, there is also an option to filter the data based on source by using the buttons to the left of the chart.

In the three following tabs, i.e. Twitter, Facebook and Blogs, a user can drill down to view more detailed information regarding each source. The screen shot in Appendix IV shows the Twitter tab which includes three different charts, visualizing the three different metrics chosen for this study, i.e. mentions, sentiment and influence. A line chart was chosen in order to show the change in mentions over time (see figure 4), whereas a bar chart was chosen to show the average sentiment during a specified period of time. Influence was displayed in a sorted table, including the 100 most influential tweeters together with their Twitter picture, a link to their Twitter page, their average sentiment and their Klout score. Moreover, the Facebook and Blog tabs are almost identical to the Twitter tab. The only difference is that the Facebook tab includes a bar chart showing average likes per post instead of a list of influential users.

The third and final level of analysis abstraction is provided by the Details tab, see Appendix V, where the original texts are listed in a table for each source. These tables include all retrieved messages together with information about the user who wrote it, when it was written, its calculated sentiment and the links it contains. There is also a summary table for each source summarizing the values of the metrics, e.g. average mentions per day and average sentiment. Further, the filtering options on the left enable a more detailed analysis since a user can choose which messages to look at depending on interesting findings from previous tabs.

5.2 Information Quality Assessment

In this study the quality of the data retrieved from the different sources was evaluated with inspiration from the framework by Wang and Strong [59], which was previously described in section 3.5. The main focus lied on assessing the contextual and representational quality, i.e. the relevance and interpretability of the content, whereas intrinsic quality issues only were briefly noted. Since the accessibility aspect refers to the properties of the information system, its assessment would in this case imply the evaluation of QlikView and its features. However, the main purpose of the study is to investigate if information from social media can be of any relevance to release planning. Hence, the accessibility aspect was considered to lie outside the scope of this study.

In terms of content relevance, this was investigated by manually reading through 30 messages for each subject and for each social media platform, i.e. a total of 270 messages. Another aspect that was considered when assessing the contextual information quality, was whether or not the amount

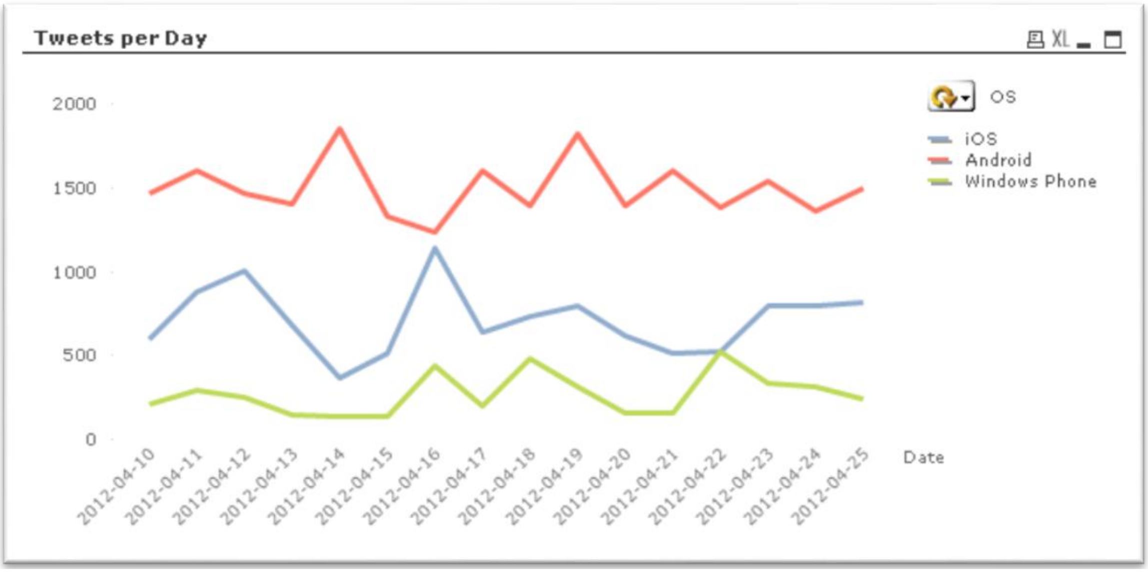


Figure 4 - PoC application screen shot: Twitter chart of mentions

of provided information was sufficient. This was investigated by calculating the volume of the fetched data from each of the chosen social media platforms. Further, the different subjects were also compared in order to draw conclusions about which one is mostly talked about and therefore maybe more appropriate to use.

Further, the interpretability of the content was investigated by comparing the accuracy of two different sentiment algorithms, Alchemy[19] and Repustate. This was done by manually scanning through a hundred messages, but in this case only for one search term and only for Twitter and Facebook. Each message was first scored according to personal judgment and then compared with the scores acquired by the algorithms. Moreover, other interpretability issues discovered during the scoring process were noted.

5.3 Evaluation

The PoC application was evaluated in two companies within the telecom industry, Ericsson and Company X. The main case is the one of Ericsson, which provides a proper software release planning perspective to the evaluation. However, the release planning in this case is mainly driven by a business-to-business perspective. The evaluation at Company X, which operates in a market-driven context but in terms of distribution rather than software development, is therefore seen as a good auxiliary case to cover the market-driven perspective of planning.

5.3.1 Case 1 - Ericsson

Ericsson[20] is an international telecommunication company and currently the world's largest vendor of telecommunication equipment [13]. Among other products, they develop mobile network data switches, of which mobile network operators are the main customers. The software release planning for this product is hence mainly driven by requirements from network operators, but also by regulations, standardization and internal requests. As in many cases of software release planning, one challenge is to determine which requirements provide the highest business value, e.g. how to best help customers to become more profitable. In order to achieve this it could be beneficial to be informed about current opinions of network end users, i.e. the customers of network operators, even though the release planning is not market-driven in a traditional sense.

For evaluation of the PoC application, a semi-structured group interview was conducted at Ericsson. The participants of this interview had the specific roles of Product Manager and Sales Support Manager at the department developing mobile network data switches. Both of these roles had deep knowledge about the development of the product as well as customer requirements and market situation in terms of competitors. They also had common knowledge about popular social media platforms as Twitter and Facebook, but none of them were previously familiar with social media mining or the QlikView platform.

The goal of the interview was to evaluate our hypothesis, i.e. that a BI tool such as the PoC application potentially could support the process of software release planning by providing valuable insights into consumer opinions. The agenda of the meeting had the following structure. First, a presentation of our hypothesis and relevant background was held to introduce the participants to the purpose of the interview. The presentation was followed by a live demonstration of the PoC application, showing the intended workflow and graphical interface. The remaining part of the meeting was spent asking open-ended questions about software release planning, current decision

[19] Alchemy: http://www.alchemyapi.com/api/register.html
[20] Ericsson: http://www.ericsson.com/

support and the use of an application similar to the demonstrated PoC as a support tool for software release planning at Ericsson. For a complete list of interview questions, see Appendix IX. In addition, the workflow of the application in terms of a potential Ericsson specific scenario as well as a general software release planning scenario was discussed with the participants in retrospect.

5.3.2 Case 2 – Company X

Company X is a distributor of wireless telecommunication devices, e.g. mobile phones, with over 25 000 business-to-business customers globally. One of their challenges is to keep up-to-date with market trends in order enable fast delivery of demanded products. Keeping the right products on the shelves at the right time is crucial, and prioritization within procurement of products can be compared to the challenges of software release planning as described in section 5.1. Differences in this comparison however include types of flexibility and time aspect. A product planner often has a longer time perspective than a purchase planner because of the time it takes to actually develop products. This can be seen as a benefit for the purchase planner because of the possibility to more rapidly respond to changes in market demand. However, a distributing company is still dependent on the range of products available on the market, hence in this sense product developing companies have more control of evolution of products.

The evaluation of the application in Company X was conducted using a semi-structured group interview, where the roles of the participants are presented in table 6. The participants' experience level regarding the QlikView platform ranged between intermediate and expert.

Roles	Input Perspective
Purchase Planner	Similar needs for decision support as a product planner.
Sales Manager and Market Coordinator	Good insight into market trends and demand.
IT Manager and Business Analyst	Deep understanding for technical solution and BI.

Table 6 - Roles of interview participants

The meeting was structured similarly to case 1, starting with a presentation of problem background and our hypothesis followed by a live demonstration of the product and ending with a discussion led by open-ended questions. Since general questions about the company and different roles interested in the PoC (e.g. purchase planner) were already asked during requirements elicitation interviews at Company X, the questions in the group interview mainly regarded the usefulness of the PoC as summarized in Appendix X.

The participants at Company X also had the opportunity to try the application during one week before the evaluation meeting in order to gain insights and raise potential questions or unclarities. Because of restricted QlikView licenses, this opportunity was not possible in case 1.

5.3.3 Analysis

The group interviews were recorded in both of the above cases for the purpose of further analysis, during which similarities and differences between the cases were identified. Statements and parts of the discussions which could confirm or contradict our hypothesis were recognized, as well as new insights and suggestions about future improvements based on the PoC application.

5.4 Validity

In order to give a better understanding of the outcomes of this study, the validity of the chosen research design is discussed from two perspectives: external and conclusion validity.

5.4.1 External Validity

Decisions made during the creation of the research design, such as choice of metrics, subjects, sources, tools and language, might all affect the obtained results of this study. Firstly, the case study is conducted within the telecommunications industry and, therefore, the result might be especially interesting for roles in that specific business area. Some of the chosen subjects, however, such as operating systems or unreleased products can be considered to be unrepresentative for the general consumer market, since they might only be discussed in more technical circles. Also, investigating only three different search terms per subject limits the extensiveness of the conducted research since other potential items, within the same area, are left unexplored.

The limitation of only gathering English messages implies that a global view of buzz on social media could not be obtained. Also, the only way to fetch data in a specific language (through the Facebook and Twitter APIs) is to search for all messages written by users who have configured their *page language* setting to English. This means that messages in other languages than English or messages that were posted by users in other countries than the US or Great Britain might end up in the results, hence the gathered content does not give an accurate representation of the English market either.

Moreover, other relevant online sources (than those used in this study) include for instance product rating sites, forums and YouTube. Other metrics, such as interconnections between users or the amount of user activity on product fan pages on Facebook, could also have been considered but lies outside the scope of this study. Additionally, the chosen tools for fetching and visualizing data, i.e. QlikView and QVSource, have not been evaluated to be the optimal solutions for this study. Hence, other solutions could also be appropriate.

5.4.2 Conclusion Validity

As mentioned in the delimitations chapter, an ideal case for the evaluation of our hypothesis could not be obtained. Since the application was evaluated in two complementary contexts, the results and conclusions of this study might not mirror the results and conclusions that could have been obtained if the evaluation had occurred together with a product planner in a market-driven software development context.

Furthermore, since Company X was involved from the beginning of the development, they had greater insight into the possibilities of the PoC during the evaluation group interview than Ericsson had during their interview. The reference group at Company X also had the possibility to try out the application during one week before the evaluation interview, but the main part of the group did not have time to do this. Also, the conducted interviews were limited in terms of time which implied that questions had to be prioritized accordingly. In conclusion, more extensive feedback could have been received from both cases under different circumstances.

Regarding the gathered data sample, the completeness aspect of the contextual information quality could not be guaranteed since all the search terms used were very specific e.g. “Samsung Galaxy S3”. This might lead to misleading information (e.g. incomplete number of mentions) since a mobile phone could for instance be referred to as, “the next generation of Samsung Galaxy”.

6 Results

The developed PoC application was, as presented in previous chapter, used to conduct a case study with the purpose of evaluating our hypothesis. This chapter presents the results and interpretations of the case study, both in terms of information quality assessment for gathered data and evaluation within the companies.

6.1 Information Quality Assessment

The quality of the gathered data was assessed according to the three different information quality categories described in section 3.5. When manually analyzing the content, several intrinsic issues appeared affecting the accuracy of the amount of mentions that a specific subject gained. Such inaccuracies mainly included spam and commercial messages, which should not count as consumer opinions. An example is the following commercial tweet, which talks about an offer by Vodafone:

*“** VODAFONE SUPERDEAL GRATIS SAMSUNG GALAXY Y ** http://t.co/ZSTELcdu”*

In figure 5, results for the contextual information quality, i.e. relevance of the content, are presented. Between the three sources, Twitter contained the most relevant content attaining a value of 88% whereas Facebook and Blogs received lower values.

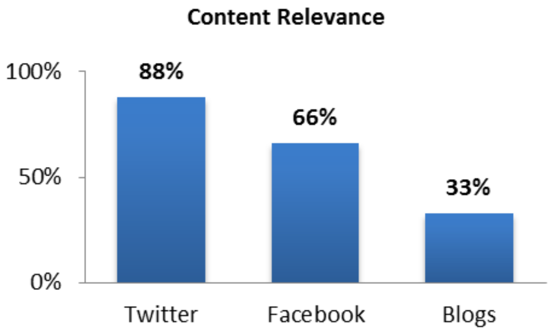


Figure 5 - Result of contextual assessment

Since this information should serve as a support for decision making, values closer to a 100% are preferable in order to increase the confidence of information consumers. The content relevance could be improved through various forms of search optimization. For instance, when searching for posts and tweets containing the operator name “vodafone”, many results included messages that referred to a music contest called Vodafone Icon. One way to optimize the results in this case could be to search for all messages including the word “vodafone” and excluding the word “icon”. This type of functionality is provided by all the chosen social media APIs and could be useful for filtering out irrelevant noise.

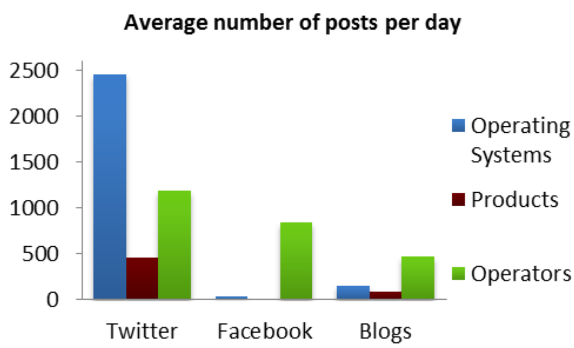


Figure 6 - Volume of gathered data

Regarding the volume of the content provided by each source, it is clear that the highest amount of data was retrieved from Twitter, as shown in figure 6. The lower value for Facebook could be explained by the fact that they enable their users to limit access to their information, in contrast to Twitter where all data is publicly available. Further, operating systems have the highest number of mentions on Twitter, whereas on Facebook it is operators that attain the highest amount of mentions. Both subjects together are more talked about on the compared social media sources than upcoming products.

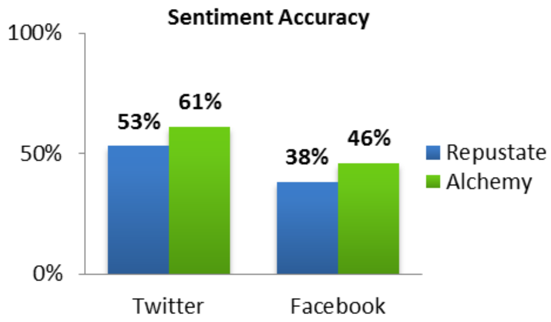


Figure 7 - Result of representational assessment

Figure 7 presents the results obtained from the assessment of the representational information quality and shows that the accuracy of the two compared sentiment algorithms is fairly low. Although Alchemy seems to be slightly more accurate than Repustate, this is probably because they employ different scoring scales where Repustate uses discrete values, i.e. -1, 0 or 1, and Alchemy uses a continuous scale ranging from -1 to 1. Moreover, since these algorithms are commercial it is hard to find information about how they actually work, which in turn makes it hard to draw conclusions about why they were more accurate for tweets than Facebook posts. However, with the assumption that these algorithms are based on the approaches described in 3.4, it can be argued that shorter posts are easier to analyze than longer posts. This because they probably contain fewer adjectives (sometimes only one), which simplifies the sentiment analysis process.

The challenges that come with interpreting social media messages conform with current research about the subject [9, 61]. Below are some examples from our data sample that illustrate these difficulties:

"Got my new iPhone ... I was sick without it"

This post is obviously positive for iPhone. However, since the word "sick" is included, the algorithms classify it as negative. Similar examples include messages with irony and slang.

"Goodbye blackberry era. It's been swell. iPhone, your time is now"

This example can be judged as either positive or negative depending on whether it is analyzed from iPhone's or Blackberry's perspective. The algorithms do, however, not understand perspective and classify the message as being positive or negative for both iPhone and Blackberry.

"Forgot my iPhone charger and I'm on 16% #WorstFeeling"

This tweet is a typical example of the complexity encountered, even by humans, when evaluating sentiment. For instance, it is unclear whether this message refers to iPhone's bad battery time or if the person who wrote it simply has a bad memory.

To summarize, data from social media includes noise and can be irrelevant if the search approach is not optimized. Further, Twitter serves as a rich data source compared to Facebook and Blogs and contains more relevant data. Finally, the sentiment algorithms used have low accuracy, which is not surprising because of the already acknowledged challenges associated with interpreting social media data.

6.2 Evaluation

The PoC application was evaluated in group interviews at two different companies as described in section 5.3. Below is a summary of each interview together with a few example quotes as well as an interpreting analysis of the results.

6.2.1 Ericsson

The interview participants at Ericsson, who in general were positively overtaken by the possibility to analyze content on social media, received our hypothesis with optimism and empathized the potential of the PoC tool. Specific applications were, however, not completely obvious in the context of release planning for their product, i.e. mobile network data switches. A curiosity of getting to know about trends in consumer opinions on social media was continuously weighed against the fact that development of the product is primarily driven by customers (i.e. network operators) and not end-users.

The widely accepted view of release planning as a complex task was aligned with the interviewees' experience, compiling requirements from different sources and prioritizing them to maximize business value was considered challenging. The majority of the requirements originate from customers, standardization and regulations, where the two last-named are fundamental for selling the product. A lot of effort is also placed on understanding the needs of the network operators since it is of course important for Ericsson to understand and support the business models of its customers, the operators.

As described in 3.1, it can be argued that Ericsson in extension is driven by market demand (i.e. by end-user needs) since network operators are dependent on what consumers are willing to pay for.

The interview made clear, however, that although discussions about end-user behavior are common at the department, most focus is placed on the network operators. Also, a distinction between consumer behavior and consumer opinions should be made. Consumer statistical usage patterns and network key performance indicators (KPIs) are continuously measured and collected through the network data switches and other network elements by the network operators. Besides measuring the performance of specific network operators, this information can also be used by Ericsson to optimize network performance. Consumer opinions, however, cannot be captured by only logging network traffic and one of the interviewees pointed out this gap in current decision support:

"The operators gather and measure [network key performance indicators], but that does not show how the end-users are experiencing the service"

The interviewees did mention Ericsson ConsumerLab[21] as a source of market research on consumer trends, but also commented that these annual reports are often describing high-level trends useful for long-term strategic decisions rather than for detailed release planning decisions about the their product. A BI tool as the PoC could possibly provide more specific and up-to-date information about consumer opinions relevant to the product planners of the network data switch. One example scenario brought up by the interviewees was to analyze how the network performance is perceived by users at different geographical locations, which is not necessarily aligned with the measured network performance. Although geographical attributes are not supported in the PoC, a big potential of the tool mentioned by the interview participants was to provide more objective insights into consumer opinions:

"If it becomes widely known that everybody complain about a certain operator, then [Ericsson] wants to know"

On the contrary, the interviewees also stated that it would probably be difficult to find anything based on consumer opinions that would actually change the way the network data switch is developed, i.e. which features to be included in upcoming releases. If it is found that consumers are unsatisfied with a certain operator who utilizes Ericsson switches, this would probably be communicated to the customers together with a recommendation to upgrade their network setup or expand capacity, i.e. acquire more comprehensive equipment from Ericsson. Although this business application would support sales, it is not supporting the hypothesis of social media mining being valuable input for software release planning. The interviewees also stressed that the main user demands, e.g. high mobile connectivity and possibility to perform common tasks and use popular services on the Internet, are already known and that users in this sense are very similar and predictable. More interesting would be to analyze differences between various devices and operating systems because they operate in different ways to connect to the network which affects the performance. It was suggested that analyzing the experienced performance for different devices and operating systems could help identify possible network optimizations.

Over all, being informed about upcoming requirements and future demands on mobile networks are of course important for Ericsson as it is part of their brand to be trustworthy and able to provide expert advice to their customers. But when discussing potential applications of mining consumer opinions through social media, the interviewees indicated repeatedly that this is perhaps more relevant for the operators directly (and not through Ericsson). It was pointed out that most network operators are already doing extensive market research to get insight into consumer opinions regarding their networks and that the use of a tool similar to the PoC could be a cheaper, complementary support for them to gather this kind of insights.

In conclusion, no obvious application of social media mining as input to release planning of the mobile network data switches was found during the interview. One reason for this could be that the subject had not been discussed by the participant prior to the interview. One of the interviewees expressed that:

"I didn't even know that it was possible to extract this kind of information [from social media]"

As mentioned in the beginning of this section, the interviewees were positive to the idea of social media mining as input to software release planning in general. In the specific network data switch case, however, the participant wanted more time to think about potential applications. The question whether Ericsson should perform this kind of consumer opinion analysis even though not directly market-driven was also considered to need further investigation.

Later, in retrospect to the group interview, a potential scenario was discussed with one of the interviewees. During the interview, most focus was placed on analyzing negative buzz on social media to identify problem areas, but afterwards the product planner at Ericsson suggested to also focus on positive buzz around operators who utilize Ericsson

[21] Ericsson ConsumerLab: http://www.ericsson.com/thinkingahead/consumerlab

switches in order to strengthen competitive differentiation. The suggestion could be supported by empirical evidence provided by Karlsson and Ryan [27], showing that intrinsic customer satisfaction focus (i.e. customer satisfaction is seen as the absence of dissatisfaction) is less efficient for improving performance than extrinsic customer focus (i.e. finding new ways to increase the positive aspects of customer experience). Based on this theory and the steps of the intended PoC workflow described in figure 2, the following example scenario was suggested and confirmed to be relevant for Ericsson:

1. For the upcoming product release of the Ericsson mobile network data switch, the product planner wants to strengthen the product differentiation by including features which can reinforce positive user experience.

2. The product planner chooses to analyze buzz around three different telecom operators, where two of the operators (X and Y) are using Ericsson switches. The analysis shows that there lately has been slightly more buzz around the operator (Z) which is not using Ericsson switches relative to the others.

3. The buzz for operator Z is overall negative, while for operator X and Y it is relatively positive.

4. The influence of the content authors is fairly equal in this case.

5. Digging into buzz details, the product planner finds that the positive buzz around operator X and Y is mainly due to favorable comments about fast connection. At the same time, the negative buzz around operator Z is partly because of complaints about slow connection speed.

6. For the next release, in addition to fundamental requirements needed to sell the product, the product planner will therefore prioritize optimization of connection speed in order to strengthen the competitive differentiation.

In addition to the confirmation of above scenario to be relevant for Ericsson, a fictive scenario for software release planning in a more market-driven context was also presented to the product planner at Ericsson. Based on his knowledge within release planning he considered this scenario to be completely logical and realistic, which supports the credibility of our hypothesis. The scenario, structured according to the intended PoC workflow, was presented as this:

1. For the upcoming product release of the mobile operating system X, the responsible product planner wants to include features which somehow can improve the competitive advantage.

2. The product planner chooses to analyze buzz around operating system X as well as around competitors' operating systems. Products that are using these different operating systems are also analyzed. It is found that the buzz around operating system X is slightly lower than for other operating systems.

3. Additionally, the buzz around operating system X is also more negative than for other operating systems.

4. There is a group of high influence content authors in the area of mobile operating systems and related products.

5. By digging in to buzz details, it is discovered that several of the high influence content authors are recommending products with other operative systems than X because of X's relatively bad battery time.

6. For next product release of operating system X, the product planner therefore chooses to prioritize requirements for optimization of battery time.

6.2.2 Company X

In comparison with the Ericsson case, the interviewees at Company X were also positive towards the use of a tool as the PoC to analyze market trends through social media content. The participants considered the extracted information to be a valuable complement to existing decision basis for procurement decisions. Even though content relevance and sentiment accuracy is not 100%, the information would contribute to a "gut feeling" which is useful when making decisions.

The information about consumer opinions through a BI tool was also considered to contribute to smarter discussions externally. For instance it was suggested that this kind of information could be presented as a service to resellers. It would also be beneficial to have the information at negotiation with suppliers (e.g. manufacturers) to be able to confirm or argue against the suppliers' claims about product popularity among consumers. The interviewees also agreed that the tool would probably be interesting for product planners at a manufacturing company.

Regarding specific subjects to search for and analyze, the interview participants found it difficult to say what would be most interesting to look at, e.g. products, operating systems etc. If the tool was to be used at Company X, they would like a base of some different subjects to start with and analyze over a period of time in order to later expand or adjust the selection based on gained experience. That subjects to search for are easily changed in the PoC was considered necessary, even though a change in subject selection would lead to decreased continuity in data history.

One essential type of information pointed out to be missing in the PoC was geographical attributes of the social media content. Without being able to group the content according to location, it would be difficult to make procurement decisions based on it because of big differences between geographical markets. The purchase planner participating in the interview stressed that:

"Information about regions is crucial in order to use this in a proactive way"

For instance sentiment was considered less important than geographical location, especially considering the low sentiment accuracy and long processing times. One suggestion was to only process sentiment calculations for a limited sample of gathered data to get a hint about content attitude without spending too many resources. Learning about and optimizing the specific words to search for was also considered more important than sentiment analysis.

The graphical interface was considered well-structured and fulfilling the need for analysis at different levels of detail since it is important to both get a quick overview and also be able to drill down the information with great flexibility. It was summarized by one interviewee that:

"It is very easy to get an overview [of the information]"

6.3 Interpretation of Results

The result of the evaluation interviews, summarized in previous section, varied somewhat between the two different cases. The interview participants at Ericsson were, although positive towards the hypothesis, rather unsure about specific applications for the release planning of their product. The interviewees at Company X on the other hand suggested several possible fields of application, e.g. valuable input to procurement decisions and argument basis for negotiations with suppliers. The more convinced and accepting attitude of Company X could to some extent be a result of earlier involvement and more time to process such ideas than in the case of Ericsson. Another credible reason for this difference is that mining consumer opinions is more naturally connected to a market-driven context.

A mutual desire in both evaluation cases was to get additional information about geographical locations connected to the social media content. Interviewees at Company X highlighted the importance of grouping content based on separate geographical markets while the Ericsson interviewees elaborated on analyzing experienced network performance based on location. It would consequently be interesting to expand a BI tool such as the PoC to include this kind of information. A smaller investigation by the authors shows that geographical information for posts appears possible to gather from at least the Twitter and Facebook platforms (via the APIs used in this study). The completeness and extent of this information is, however, unknown by the authors and would need further investigation.

As found from the evaluation results at Company X, sentiment was considered to be less important than the desired addition of geographical locations, especially because of the algorithms' low accuracy and long processing times. One potential solution to reduce processing effort could be to only analyze a sample of the retrieved messages or to limit the data sample to only include messages posted by users with high Klout scores. However, the correlation between Klout scores and content relevance is not certain and would need to be investigated. Also, there is always a risk that high-influence users are subjects to product placement, thus, making the information they share biased.

In order to attain more relevant data, the search approach could be optimized and a learning period is probably needed in order to investigate what is appropriate to search for. Although search improvements can be achieved, there will, however, always be uncertainties and lacking information quality because of the unstructured nature of user-generated content. Considering how to visualize the altitude of uncertainty for presented information in the support tool would therefore be relevant in order to give more confidence to the decision makers.

It is important to carefully explain to potential users that a BI application based on unstructured, user-generated data should not be treated as an thoroughly accurate source of information. As described by Carlshamre [7], there is a difference between *prosthetic systems*, which generates ready-made solutions, and *instrumental systems*, which should be considered as one of several sources of information and which need to be complemented with human intelligence. The latter is seconded by Ruhe and Saliu [49], suggesting that computational intelligence needs to be combined with human intelligence in order to attain more accurate information and, thus, make better decisions. By providing access to raw data in the tool, information consumers at least have the ability to manually check the sanity of provided information.

Although the case study has provided several useful insights connected to the suggested hypothesis, the application has still not been tested in practice, i.e. the use of the tool has not been observed in a real release planning environment. Hence, it could be interesting to investigate the results of such an evaluation and to find out whether there exists any value in making information about consumer opinions directly available for product planners. Further, a comparison study, evaluating the use of the application in a strategic versus an operational level of release planning, would also be relevant to further investigate the hypothesis of this study.

7 Conclusions

The case study presented in this paper indicates potential of a BI tool for analyzing social media since the approach was conceived as a cheaper and more flexible way to find out about consumer opinions, as opposed to traditional market research. In addition to current features of the PoC application, both groups in which the tool was evaluated suggested geographical attributes as an interesting, or even crucial, type of information. The information quality of gathered data was estimated to be relatively low, but can be improved by search optimization. Additionally, it is essential to treat the information as one of several sources for input to decisions, which are ultimately made by humans. Further, the results presented in this study might be affected by the non-idealistic evaluation cases, i.e. evaluation within business-to-business software release planning and market-driven distribution separately.

Recommended future work includes evaluating the hypothesis in a pure market-driven software release planning context, potentially also investigating the suitability for strategic versus operational release planning. Also, the possibility of incorporating geographical information in a social media mining tool is recognized as an interesting field for further investigation.

8 Acknowledgements

The authors would like to thank Dr. Miroslaw Staron for his continuous feedback and valuable guidance throughout the project.

The time and engagement of the case study participants at Ericsson and Company X are also highly appreciated. A special thanks is directed to the people in the BI team at Company X, who contributed with their support, expertise and remarkable commitment.

Additionally, the authors are grateful to QVSource, Twingly and Repustate for providing the opportunity to use their products. A special acknowledgement to Chris Brain at QVSource for his extensive support and complaisant approach.

References

[1] Abran, A. and Moore, J.W. (2004) “SWEBOK - Guide to the Software Engineering Body of Knowledge”, 2004 Version, IEEE CS Professional Practices Committee.

[2] Al-Emran, A.and Pfahl, D. (2007) Operational Planning, Re-planning and Risk Analysis for Software Releases. *Product-Focused Software Process Improvement*, Springer, pp. 315-329.

[3] *Ask Your Target Market Blog* (2012) http://aytm.com/blog/research-junction/branding-and-how-it-works-in-the-social-media-age/ (03-04-2012).

[4] Asur, S. and Huberman, B. (2010) Predicting the Future with Social Media. *arXiv*:1003.5699.

[5] Begel, A., DeLine, R. and Zimmermann, T. (2010) Social media for software engineering. *Proceedings of the FSE/SDP workshop on Future of software engineering research*, pp. 33-38.

[6] Bonchi, F. et al. (2011) Social Network Analysis and Mining for Business Applications. *ACM Transactions on Intelligent Systems and Technology* (TIST), vol. 2, no. 3, pp. 22.

[7] Carlshamre, P. (2002) Release Planning in Market-Driven Software Product Development: Provoking an Understanding. *Requirements Engineering*, vol. 7, no. 3, Springer, pp. 139-151.

[8] Chen, H. and Chau, M. (2004) Web Mining: Machine learning for Web Applications. *Annual review of information science and technology*, vol. 38, pp. 289-330.

[9] Chen, H. and Zimbra, D. (2010) AI and opinion mining. *Intelligent Systems*, IEEE, vol. 25, no. 3, pp. 74-80.

[10] Davis, A. (2005) *Just Enough Requirements Management*. New York: Dorset House Publishing, pp. 63, 65, 68 and 84.

[11] Dayal, U. et al. (2009) Data integration flows for business intelligence. *Proceedings of the 12th International Conference on Extending Database Technology*: Advances in Database Technology, pp. 1-11.

[12] Dhar, V. and Chang, E. (2007) Does Chatter Matter? the Impact of User-Generated Content on Music Sales. *NYU Working Paper No. 2451/23783*.

[13] *Ericsson*. The Company - Company Facts - Businesses. http://www.ericsson.com/thecompany/company_facts/businesses (09-05-2012).

[14] *Facebook Newsroom* (2012) Key Facts. http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 (03-04-2012).

[15] Galbraith, J. (2001) Building Organisations around the Global Customer. *Ivey Business Journal*, pp.17-26.

[16] Gassman, B. (2012) To the Point: Mining Value From Social Networks. *Gartner Business Intelligence Summit*, 6-7 Feb 2012, London, UK.

[17] Go, A., Bhayani, R., and Huang L. (2009) Twitter sentiment classification using distant supervision. *Technical report, Stanford Digital Library Technologies Project.*

[18] Hagerty, J., Sallam, R. and Richardson, J. (2012) Magic Quadrant for Business Intelligence Platforms. *Gartner Research*, id:G00225500.

[19] Han, J., Kamber, M. and Pei, J. (2012) *Data Mining Concepts and Techniques*. Third Edition. USA: Elsevier.

[20] Hao Q. et al. (2009) Generating location overviews with images and tags by mining user-generated travelogues. *Proceedings of the 17th ACM International Conference on Multimedia*. ACM, New York, pp. 801-804.

[21] Hassan, A.E. and Xie, T. (2010) Software Intelligence: The Future of Mining Software Engineering Data. *Proceedings of the FSE/SDP workshop on Future of software engineering research*, ACM, pp. 161-166.

[22] Janis, I.L., “Groupthink: Psychological Studies of Policy Decisions and Fiascoes”, 2d, rev. ed. Boston: Houghton Mifflin, 1983.

[23] Jansen, B.J. et al. (2009) Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, vol. 60, no.11, pp. 2169-2188.

[24] Jantunen, S. et al. (2011) The Challenge of Release Planning. *Software Product Management* (IWSPM); 30-31 Aug 2011, pp. 36-45.

[25] Java, A. et al. (2007) Why we twitter: Understanding microblogging usage and communities. *Proc. of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, pp. 56-65.

[26] Ji, R. et al. (2009) Mining City Landmarks from Blogs by Graph Modeling. *Proceedings of the 17th ACM international conference on multimedia*, pp. 105-114.

[27] Karlsson, J. and Ryan, K. (1997) A Cost - Value Approach for Prioritizing Requirements. *IEEE Software*, vol. 14, no. 5, pp. 67-74.

[28] Karlsson, L. et al. (2002) Challenges in Market-Driven Requirements Engineering - An Industrial Interview Study. *International Workshop on Requirements Engineering: Foundation for Software Quality*, 9-10 Sep 2002, Essen Germany.

[29] Katerattanakul, P. and Siau, K. (1999) Measuring Information Quality of Web Sites: Development of an Instrument. *Proceedings of the 20th international conference on Information Systems*, pp. 279-285.

[30] Kietzmann, J.H. et al (2011) Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, vol. 54, no. 3, pp. 241-251.

[31] Komssi, M. et al. (2011) Integrating Analysis of Customers’ Processes into Roadmapping - The Value-Creation Perspective. *Requirements Engineering Conference (RE)*, 2011 19th IEEE International, 29 Aug 2011 - 2 Sep 2011, pp. 57-66.

[32] Lappas, G. (2011) From Web Mining to Social Multimedia Mining. *International Conference on Advances in Social Networks Analysis and Mining*, 25-27 Jul 2011, pp. 336-343.

[33] Lee, S. et al. (2006) Using Patent Information for New Product Development - Keyword - Based Technology Roadmapping Approach. *Technology Management for the Global Future, 2006.* PICMET, July 2006, vol. 3, pp. 1496-1502.

[34] Lee, Y.W. et al. (2002) AIMQ: a methodology for information quality assessment. *Information & Management*, vol. 40, no. 2, pp. 133-146.

[35] Lehtola, L. et al. (2004) Requirements Prioritization Challenges in Practice. *Lecture Notes in Computer Science*, vol. 3009/2004, pp. 497-508.

[36] Marasanapalle, J. et al. (2010) Business Intelligence From Twitter For The Television Media: A Case Study. *Business*

Applications of Social Network Analysis (BASNA), 2010 IEEE International Workshop on, 15 Dec 2010, pp.1-6.

[37] *Mashable* (2010) When Are Facebook Users Most Active?. http://mashable.com/2010/10/28/facebook-activity-study/ (21-04-2012).

[38] Maylor, H. (2010) *Project Management*. Fourth Edition, Harlow: Prentice-Hall, pp.76-80 and 84-87.

[39] *NielsenWire* (2012) Friending The Social Consumer. http://blog.nielsen.com/nielsenwire/online_mobile/friending-the-social-consumer/ (03-04-2012).

[40] Park, B.K. and Song, I.Y. (2011) Toward total business intelligence incorporating structured and unstructured data. *Proceedings of the 2nd International Workshop on Business intelligencE and the WEB*, pp. 12-19.

[41] Phaal, R. et al. (2004) Technology Roadmapping - A planning framework for evolution and revolution. *Technological Forecasting and Social Change*. vol. 71, no. 1-2, pp. 5-26.

[42] *QlikView* (2011) The QlikView Product Family. http://www.qlikview.com/us/~/media/Files/resource-library/global-us/direct/datasheets/qlikview11/DS-The-QlikView-Product-Family-EN.ashx (09-05-2012).

[43] *QlikView Community* (2012) How Did We Get Here?. http://community.qlikview.com/blogs/theqlikviewblog/2012/02/28/how-did-we-get-here (03-04-2012).

[44] *Recorded Future* (2011) Big Data for the Future: Unlocking the Predictive Power of the Web, http://www.slideshare.net/RecordedFuture/big-data-for-the-future-unlocking-the-predictive-power-of-the-web (04-05-2012).

[45] Reynolds, W. N. et al, (2010) Social media and social reality. *Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on*, pp. 221-226.

[46] Ruhe, G. (2003) Software Engineering Decision Support - A New Paradigm for Learning Software Organizations. *Advances in Learning Software Organization, Lecture Notes in Computer Science*, vol. 2640, Springer-Verlag, pp. 104-115.

[47] Ruhe, G. (2003) Software Engineering Decision Support - Methodology and Applications. *Innovations in decision support systems*, vol. 3, pp. 143-174.

[48] Ruhe, G. and Momoh, J. (2005) Strategic Release Planning and Evaluation of Operational Feasibility. *Hawaii International Conference on System Sciences*.

[49] Ruhe, G. and Saliu, M.O. (2005) The art and science of software release planning. *Software, IEEE*, vol. 22, no. 6, pp. 47-53.

[50] Scott, J. (1998) Trend Report Social Network Analysis. *Sociology*, vol. 22, no. 1, pp.109-127.

[51] Shroff, G. et al. (2011) Enterprise information fusion for real-time business intelligence. *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, pp. 1-8.

[52] Stvilia, B. et al. (2007) A framework for information quality assessment. *Journal of the American Society for Information Science and Technology*, vol. 58, no. 12, pp.1720-1733.

[53] Surma, J. and Furmanek, A. (2010) Improving Marketing Response by Data Mining in Social Network. *Advances in Social Networks Analysis and Mining (ASONAM), 2010 International Conference on*, 9-11 Aug 2010, pp. 446-451.

[54] Svahnberg, M. et al. (2010) A systematic review on strategic release planning models. *Information and Software Technology*, vol. 52, no. 3, pp. 237-248.

[55] *Sysomos* (2009) Twitter Activity by Day of Week. http://www.sysomos.com/insidetwitter/#dayofweek (21-04-2012).

[56] Tumasjan, A. et al. (2010) Predicting elections with Twitter: What 140 characters reveal about political sentiment. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pp. 178-185.

[57] Turney, P.D. (2002) Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pp. 417-424.

[58] *Twitter Blog* (2012) Twitter turns six. http://blog.twitter.com/2012/03/twitter-turns-six.html (03-04-2012).

[59] Wang, R.Y. and Strong, D.M. (1996) Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5-34.

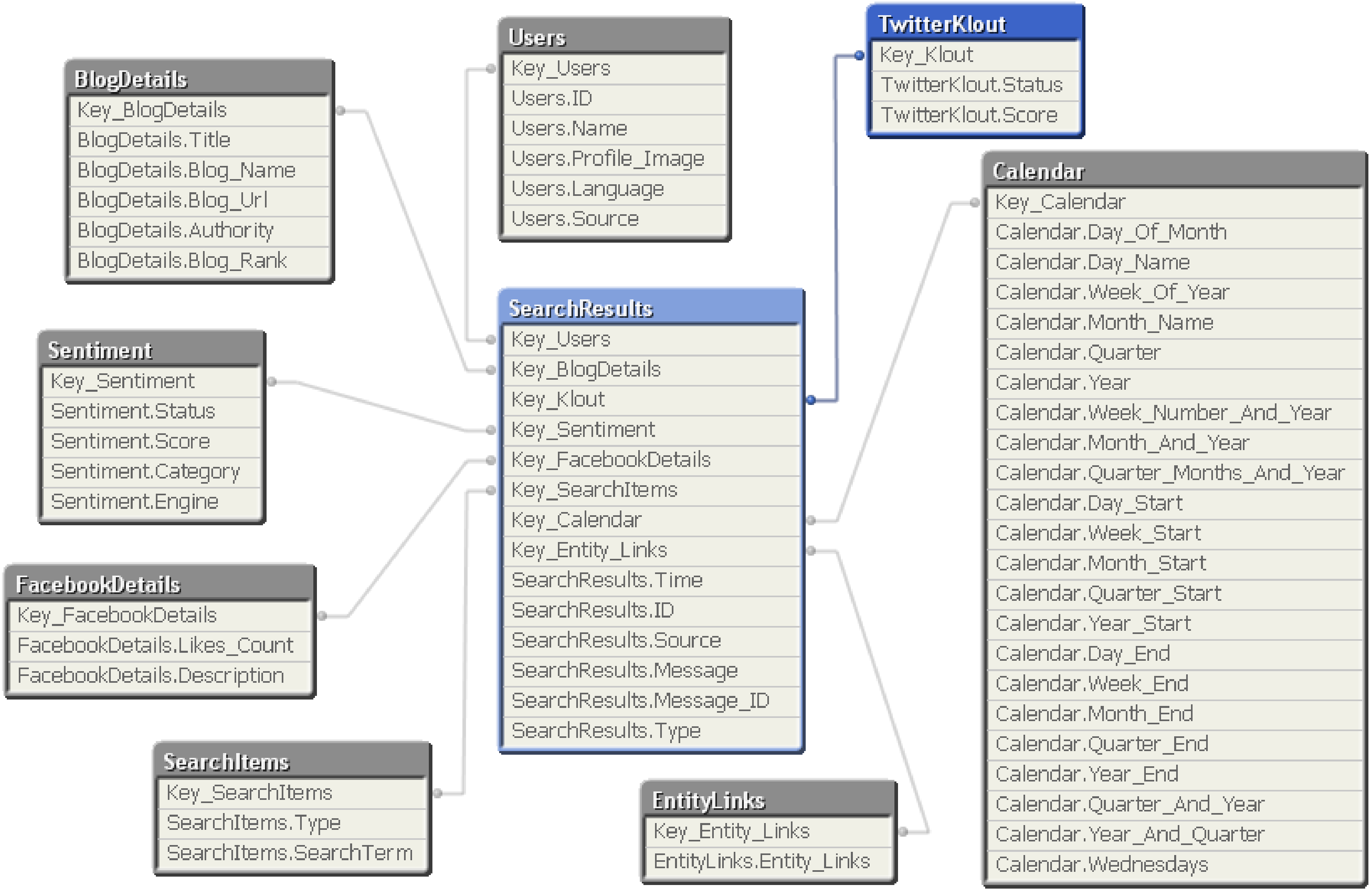
[60] Wolf, T. et al. (2009) Predicting build failures using social network analysis on developer communication. *Proceedings of the 31st International Conference on Software Engineering, IEEE Computer Society*, pp. 1-11.

[61] Yu, H. and Hatzivassiloglou, V. (2003) Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. *Proceedings of the 2003 conference on Empirical methods in natural language processing,* pp. 129-136.

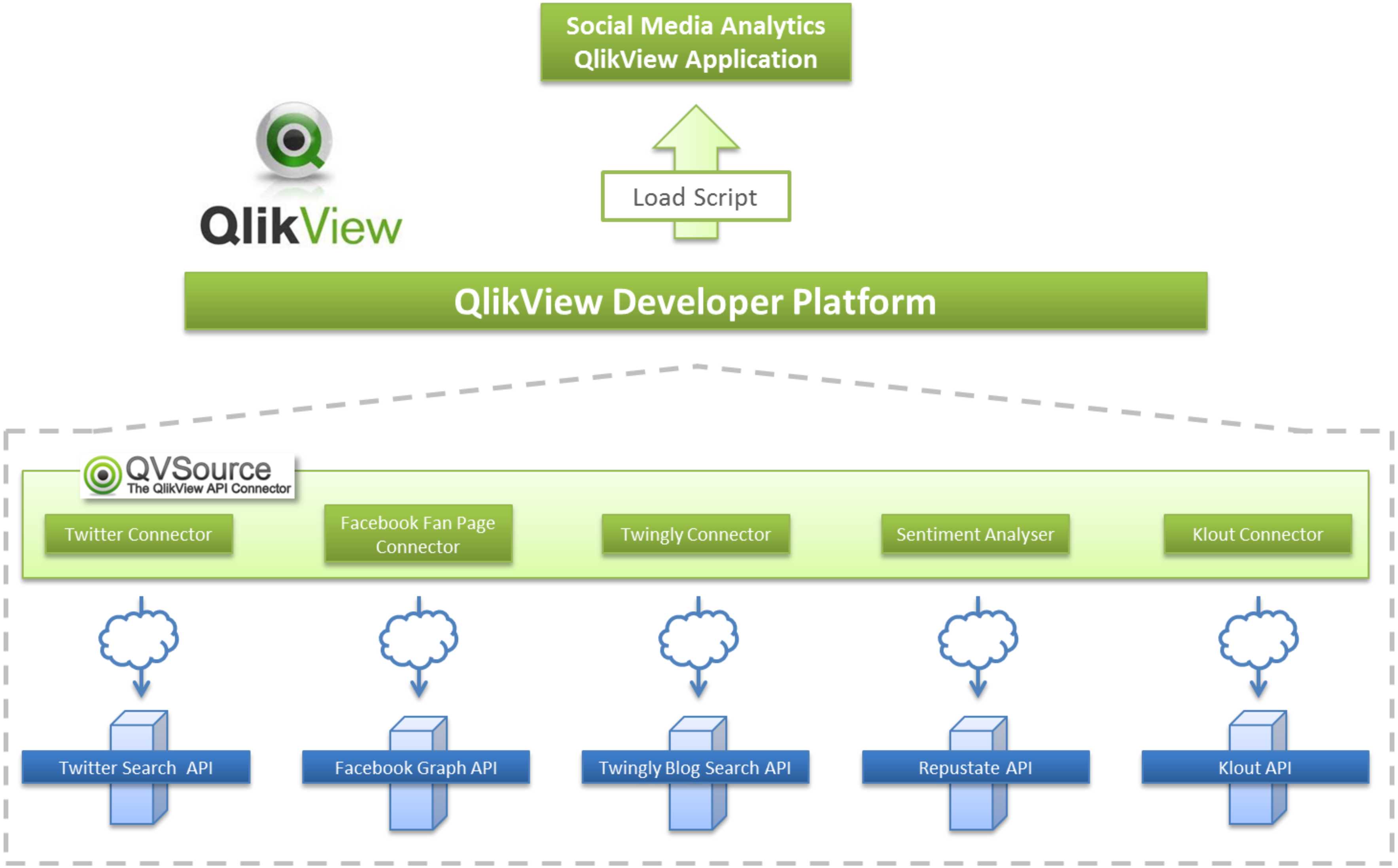
[62] Yu, S. and Kak, S. (2012) A Survey of Prediction Using Social Media. *Arxiv preprint arXiv:1203.1647*.

[63] Zheng, Y. et al. (2009) Tour the World: Building a Web-scale Landmark Recognition Engine. *Proceedings of the 2009 IEEE computer society conference on computer vision and pattern recognition*, pp. 1085-1092.

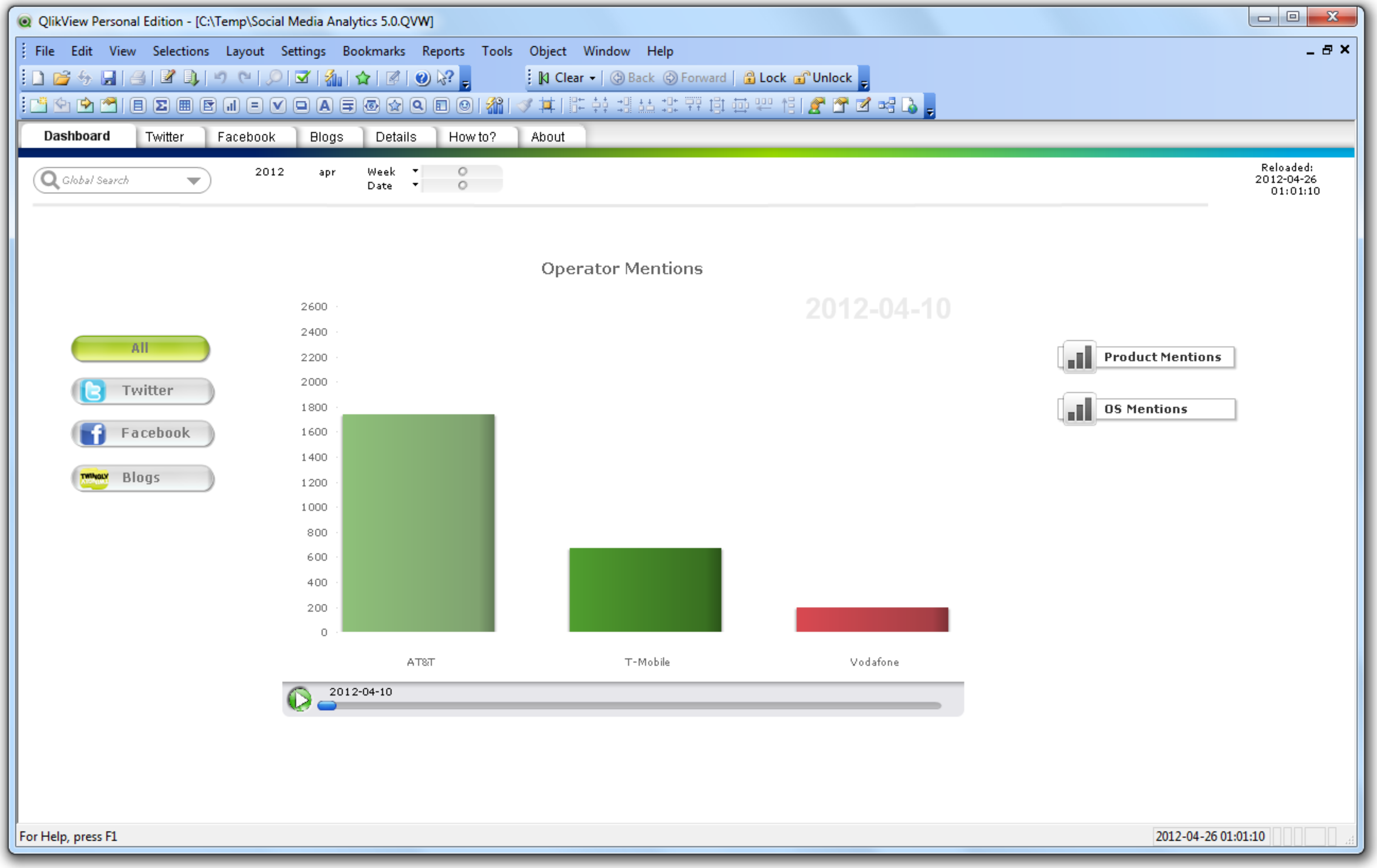
Appendix I - Data Model



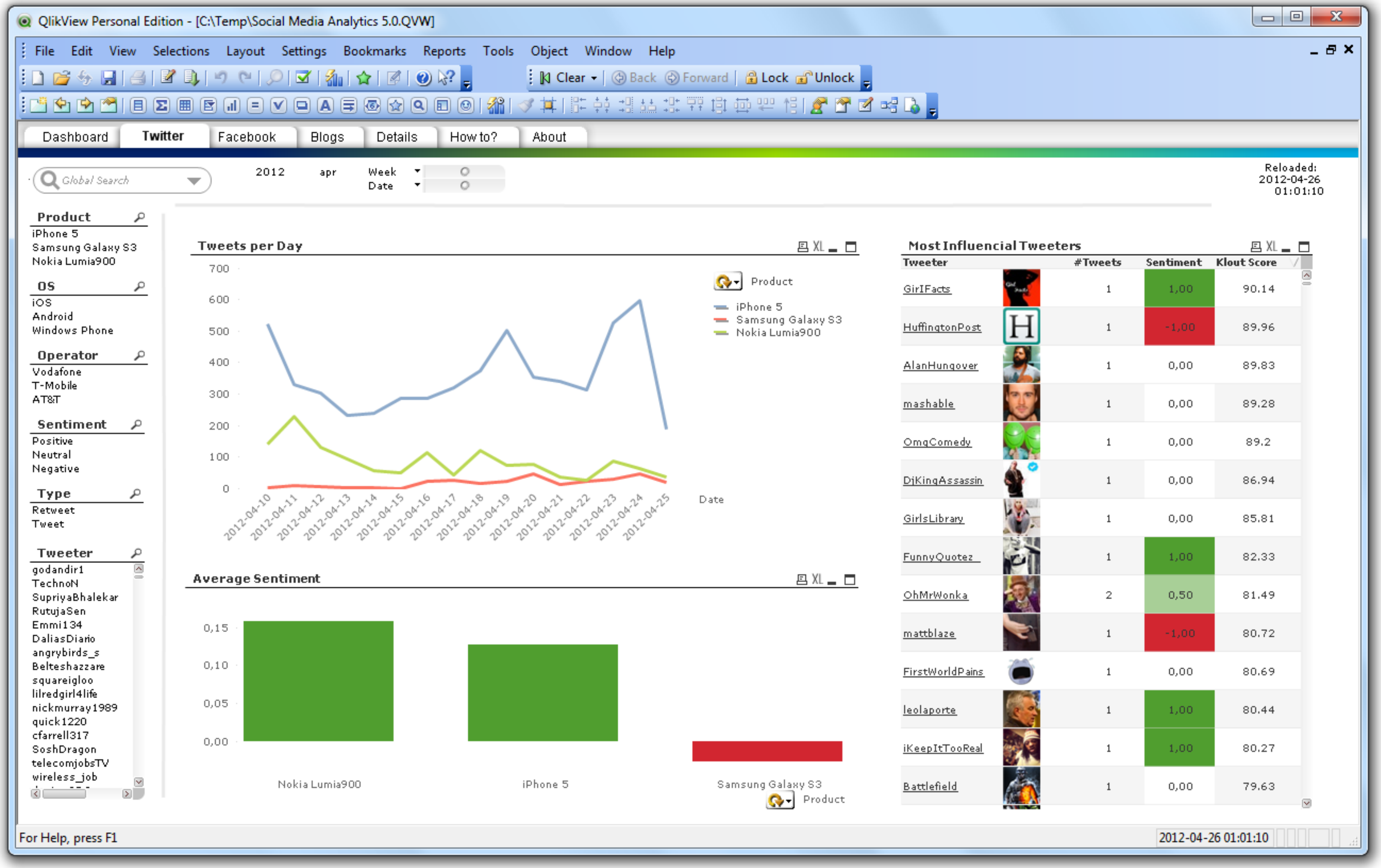
Appendix II - System Architecture



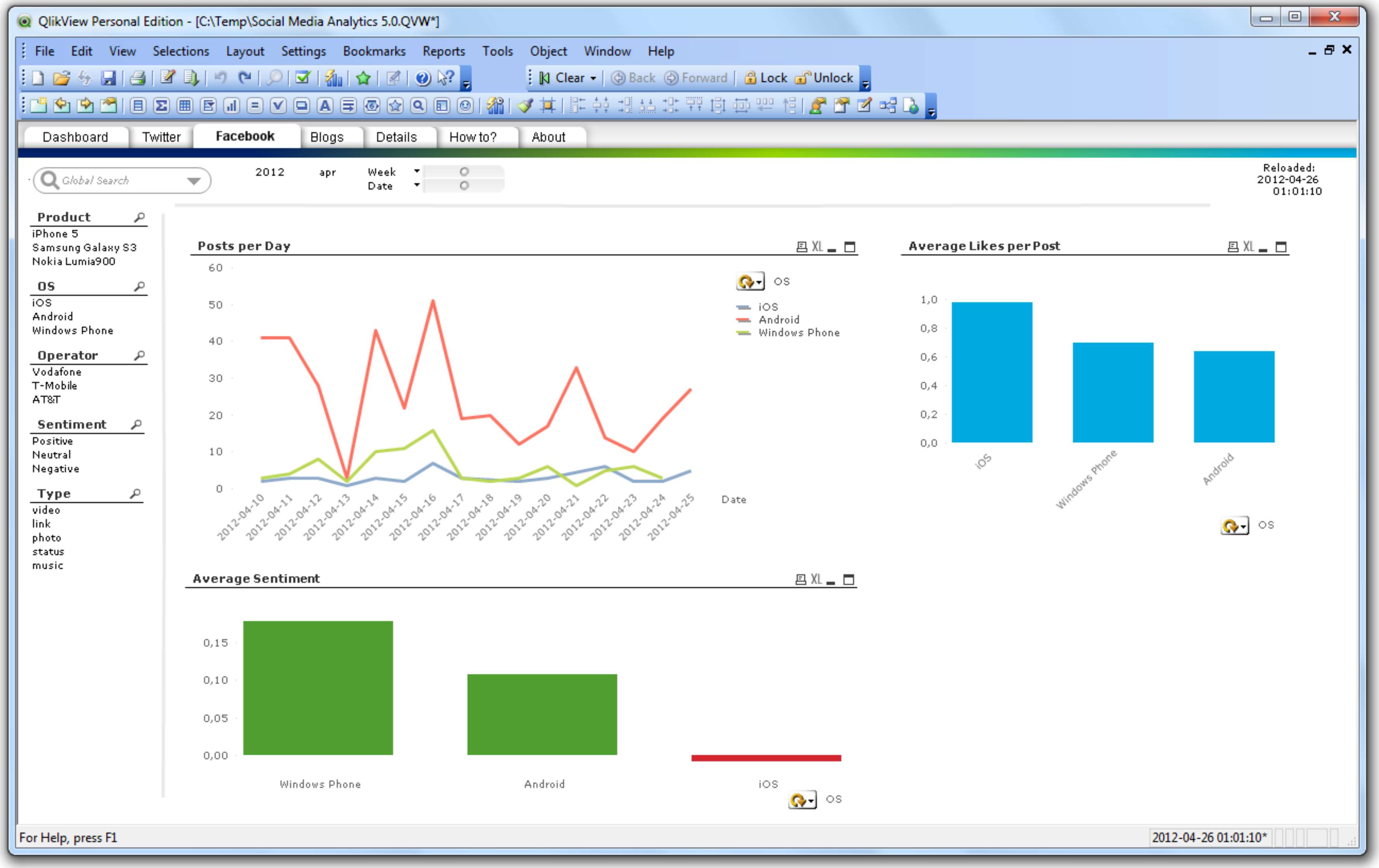
Appendix III – Screen shot from application: Dashboard



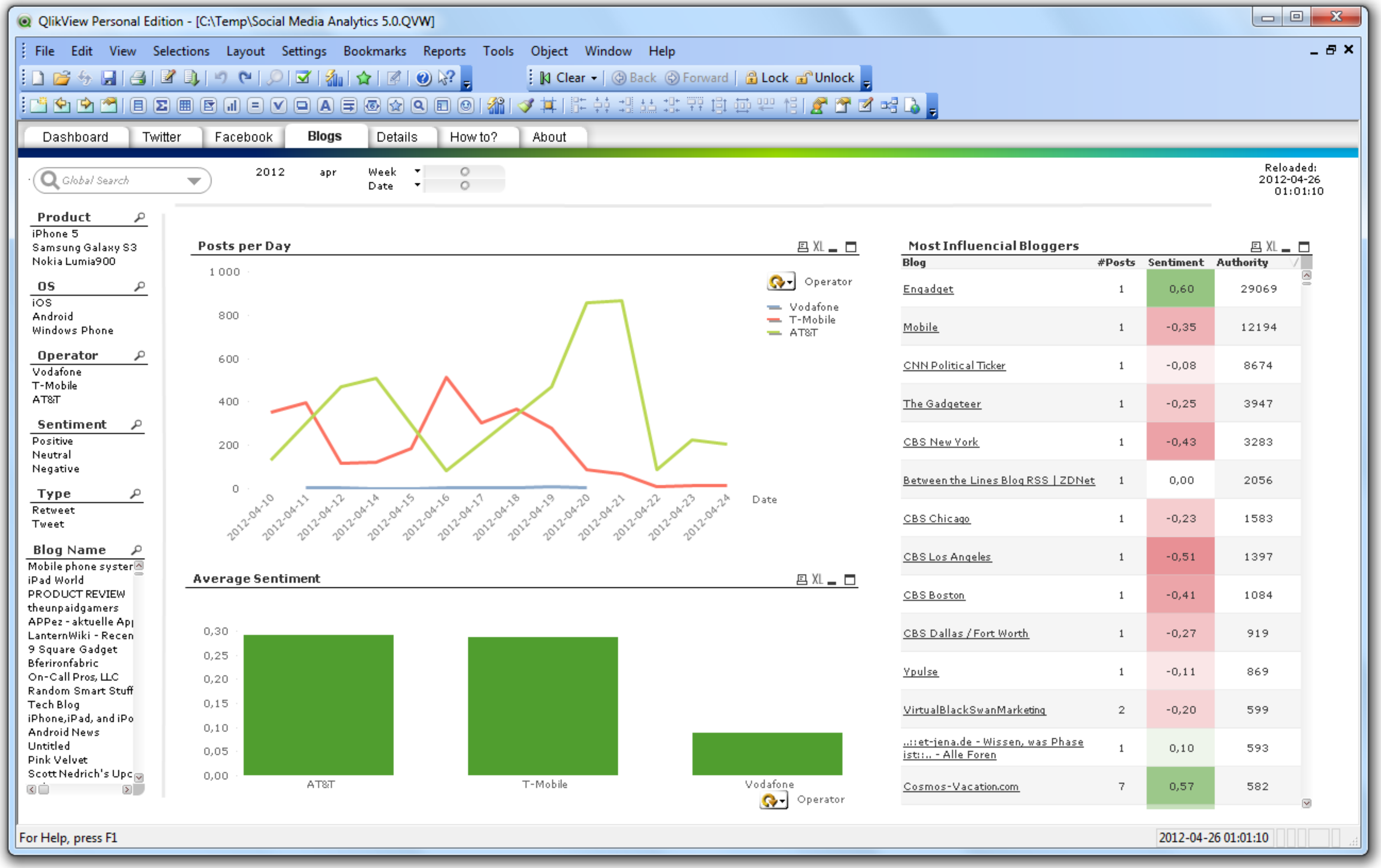
Appendix IV – Screen shot from application: Twitter



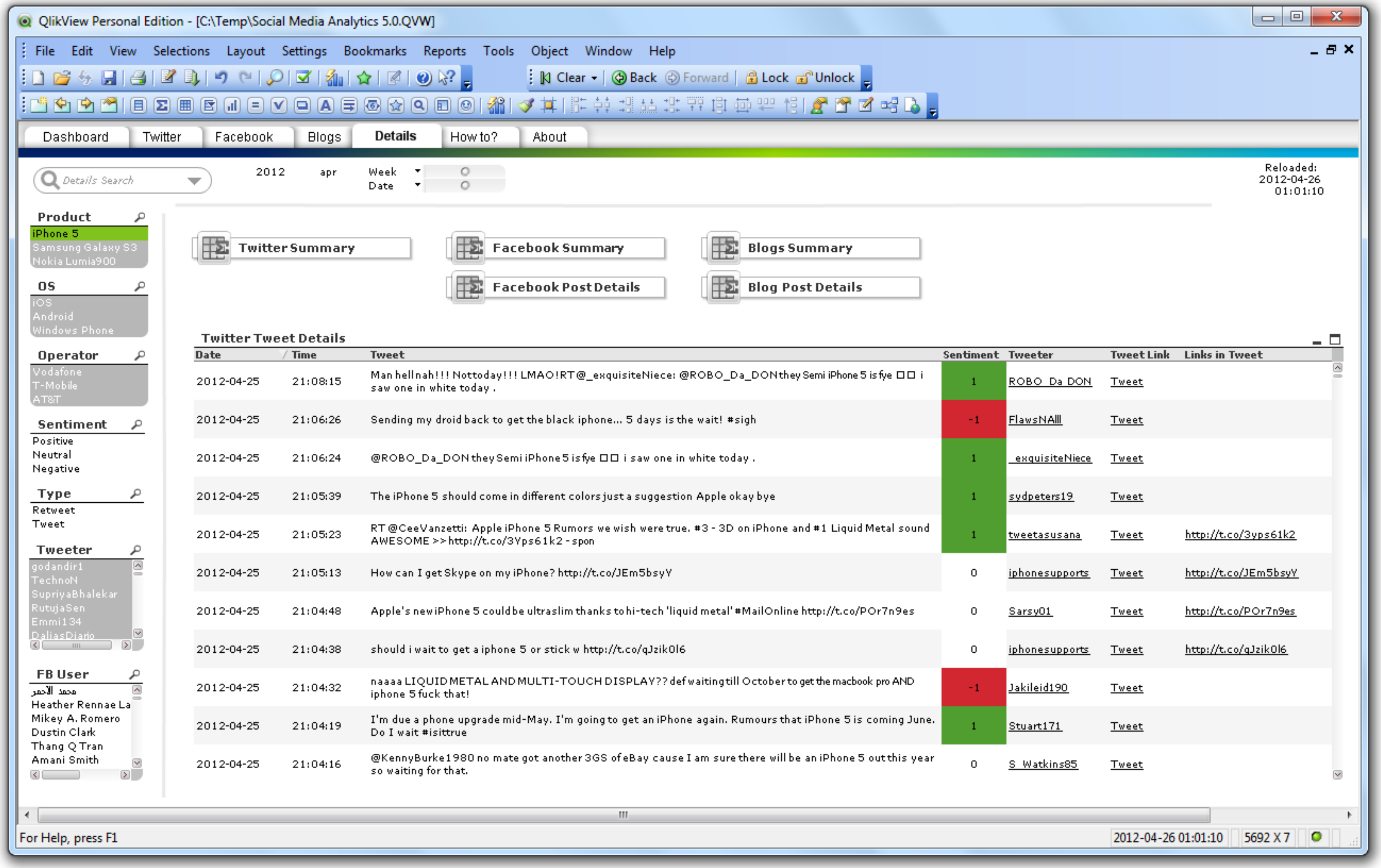
Appendix V – Screen shot from application: Facebook



Appendix VI – Screen shot from application: Blogs



Appendix VII – Screen shot from application: Details



Appendix VIII - Requirements elicitation interviews at Company X: Complete list of questions

Area	#	Question
Products	1	Can you tell us about the products that Company X is distributing?
	2	Who are your customers?
Role and Decisions	3	Can you tell us about your role at Company X?
	4	What kind of decisions is it your responsibility to take? How often?
	5	What kind of sources do you use for decision support?
	6	Regarding type of decision support, are there any common denominator between your role and other roles at Company X?
	7	Are there any types of decisions within your job that are specifically difficult to make?
	8	What do you consider the biggest challenges connected to your role?
	9	Can you give an example of a decision that you are about to make?
	10	Do you follow up on the effects of earlier decisions? How?
Social Media	11	How familiar are you with different social media platforms? (e.g. Facebook, Twitter etc.)
	12	Do you manually scan social media for input to decision making?
	13	For your role, what do think would be interesting to analyze on social media?
	14	For Company X, what do you think would be interesting to analyze on social media?
QlikView	15	What QlikView applications at Company X do you use today?
	16	How often do you use these applications?
	17	What do you think is good respectively bad with these applications?
	18	Would you prefer a highly dynamic application or more ready-made reports to support your decision making?
	19	Are there any other types of subjects you would find interesting to analyze?
Other	20	Is there anything else you think we should know?
	21	Who else do you recommend us to talk to at Company X?
	22	Can we come back to you if other questions arise?

Appendix IX – Evaluation group interview at Ericsson: Complete list of questions

Area	#	Question
Products	1	Can you tell us about the product you are working with?
	2	Who are your customers?
Release Planning	3	How are you working with release planning today?
	4	How often do you release a new version of the product?
	5	What kind of decisions do you usually take during release planning?
	6	What are the sources of requirements? (i.e. customers, regulations etc.)
	7	What kind of sources do you use for decision support?
	8	Do you use any type of BI tool for decision support? If so, what tool/tools and for what purpose?
	9	Do you conduct or use any type of market research regarding customers (i.e. mobile network operators)?
	10	Do you conduct or use any type of market research regarding end-users (i.e. mobile network users)?
Social Media	11	How familiar are you with different social media platforms? (e.g. Facebook, Twitter etc.)
	12	Do you manually scan social media content for information as input to release planning?
PoC Application	13	Do you believe that an application similar to our PoC could give valuable input to software release planning in general? If yes, what possible application(s) do you find interesting?
	14	Do you believe that an application similar to our PoC could give valuable support to the release planning in your specific context? If yes, what possible application(s) do you find interesting?
	15	Regarding the PoC, is there any particular feature/functionality that you find extra interesting?
	16	Is there any particular feature/functionality that you find irrelevant?
	17	Is there any other kind of subjects you would find interesting to analyze? (Other than products, operators and operating systems)

Appendix X – Evaluation group interview at Company X: Complete list of questions

Area	#	Question
PoC Application	1	Do you believe that an application similar to our PoC could give valuable input about market trends to the decision making at Company X? (And specifically for purchase decisions?)
	2	(If yes to question 1:) Exactly what kind of input to what types of decisions do you find interesting?
	3	How much would the input from such application affect your decisions relative to other sources of information?
	4	What types of subjects would you find interesting to analyze? (e.g. products, operating systems, operators etc.)
	5	How often do you think the subject selection needs to be updated? (e.g. for new products)
	6	Do you believe that an application similar to our PoC could be valuable to product developers?
	7	Regarding the PoC, is there any particular feature/functionality that you find extra interesting?
	8	Is there any particular feature/functionality that you find irrelevant?
	9	What would you suggest in order to improve the PoC?