LAMK Lahden ammattikorkeakoulu
Lahti University of Applied Sciences

Social media as a tool for corporate reputation management

Lahti University of Applied Sciences

Degree Programme in International Business Management

KUMLIN, MARINA: Social media as a tool for corporate reputation management.

Master's Thesis in International

Business Management 75 pages, 7 pages of appendices

Autumn 2016

ABSTRACT

_____

Inspiration comes from the fact that companies do not have control over content, timing and frequency of online conversation. A social media environment is interactive, chaotic, complex and power voice of a customer channel. The companies can't avoid or deny this phenomenon anymore. However, they can choose to utilize the opportunity to get the benefits of it. The main value of it lies in the fact that consumers or potential customers produce these public conversations.

Companies have been always looking for the tool or methodology that will transform them into more profitable, outstanding organization. Reputation of the company is an integral part of its intangible and tangible value that presents a very interesting angle for an investigation. Furthermore, the rising demand for real-time business intelligence analysis and the popularity of social media networks offer room for synergy.

The goal of the research was to identify opportunities and danger that social media may offer. The study finds that online reputation management techniques and methods may drive business excellence and create additional business value.

Twitter used as a data source for collection of stakeholders' experience and images of the case companies. Number of tweets/ retweets and positive/ negative conversation was monitored to measure the health of the company reputation. This paper aims to provide a critical look into the advantages and limitations of such the approach to monitor the reputation perceptions.

The research finds that social media has offered a host of a new platform for companies to build relationships; foster engagement and better understand their customers. The study concludes that there are many advantages and benefits for social media analysis utilisation in a business reputation management area.

Key words: sosial media, sentimental analysis, data mining, big data, internet of things, ecosystem, text mining

CONTENTS

ABBREVIATIONS

API   Application program interface

BI   Business Intelligence

DDD   Data Driven Decision

DDDM   Data Driven Decision Making

DIRE   Digital reputation

KDD   Knowledge Discovery from Data

KPI   Key Performance Indicator

ROA   Return on assets

ROI   Return on investment

RTBI   Real-time business intelligence

APPENDICES

1 INTRODUCTION

> "All truths are easy to understand once they are discovered; the point is to discover them" (Galileo Galilei)

The purpose of this thesis will be to discover how online activity may be utilized in business. The study will target acknowledgement of the opportunities that offer free available social media data. The value of the data lies in the fact that consumers or potential customers produce these data.

The worldwide accessibility to the internet is the phenomena of the modern world that is reshaping the business styles. The Internet created a technological revolution. As a result, it generated a new big data universe available for everyone. Listening to that data, making sense of it, and effectively acting on that information is essential power to any organisation. Because it is creates greater insight into how our world works. With this real time big data understanding, the companies may improve interaction with their customers. As a result, it will improves efficiency in all aspects of a business.

In fact, a consumer behavior is difficult to predict. Because of this, it is important that companies pay close attention to customers' ongoing online discussions and behavior. The companies need to understand how to use this data to win the customers and boost profits. In nowadays, important ability is to draw deep insights from oceans of data and bring them rapidly into a business decision making. A good starting point is the customer's need. The more challenging problem customer has; the more profitable result it will provide for the company. The data-driven, customer-focused companies take advantage of the insights that all this online data may deliver. To achieve this goal, the companies need to focus into the current tools and technologies to engage this big data into additional value creation for the business.

1.1 Background

> Inanimate data can never speak for themselves, and we always bring to bear some conceptual framework, either intuitive and ill-formed, or tightly and formally structured, to the task of investigation, analysis, and interpretation. (Gould, 1981, p. 166)

Big data analytics is a disruptive innovation which raises more questions than it answers. There is an urgent need for understanding this growing data, how to get deeper inside, how to approach the challenges and opportunities that big data provides.

Big data creates a wide range of opportunities and radical changes a way how we approach and think about a research process. It brings additional risk to false conclusions in the research result interpretation. Because the result of the research will not depends only on the accuracy of the result interpretation. Additionally, it is depends on tools quality that has been used in the research. It brings change at the levels of epistemology and ethics. Big data places key questions about how we should handle and deal with information. The trend of continuous big data growth brings challenges and incredible business opportunities at the same time.

The most significant source for big data is social media. CMS Wire reports in 2013, that 90 percent of the available data in the world was created over just the previous years and 80 percent of that data comes from "unstructured" sources, like social media. (CMS Wire, 2013)

Social media popularity continues to grow and is being adopted worldwide. It is tools that allow people to create and share information. The people talk to each other and share thoughts, opinions, emotions, photos, articles, useful resources and information via social media. Social media data is all about what people are doing, what they're interested in, what they're care about and what they're looking for. So, social media represents available data access to the real time survey. In top of that, social media data enables a broad range of types of analysis. The ability to collect, manage, analyze and utilize this turbulent data to gain meaningful insights will help

any business to make informed marketing and the customer centric decisions through the real time analysis.

In nowadays 'age of customer', organizations are forced to be focused on the customer experience through increasingly digital and global economy. Outside-in business strategy is an important agenda for any business. That requires a common view of what is happening in the market. Companies forced listen and respond to the customers needs to achieve the best possible customers' satisfaction. By adding value to the customer the companies can improve customer satisfaction, loyalty and advocacy.

The customer experience is one of the business differentiators that can't be eliminated. The customer's experience is a trend business focus and fundamental business driver. The short-term perspective to lower costs and increase revenue has been replaced by the need to add value to the customer by improving the customer satisfaction, loyalty and advocacy.

1.2 Research objectives and scope

> *When you understand human behavior, you improve your chances of making your business succeed. (CEO of the analytics company HiQ)*

The objective of the research will be to show the benefit of utilizing social media data mining by business in Twitter example. This shows how social media analysis can help the organisations to know their customers better and improve the success of the business. The empirical part of the research will be done by sampling and analyzing Twitter data. The outcome of the study is going to present a recommendation for utilizing social media data in company reputation management.

The data scope for this research will be large and middle size Finnish companies. The sample companies operate in different industries and have different customer relationship. The source for the social media data will be Twitter. This is permitting to gain insight into the potential differences between the Tweet's messages. Data will be collected during the period approximately half a year from 20 companies. It enables

companies to analyse the tweets post activity over time to see possible changes in the customer loyalty.

According to Smart Insight, the most popular social networks worldwide were Facebook, youTube and Twitter in 2015. (Smart Insight, 2015) For this reason, the author took Twitter as a candidate for the social media selection. Twitter is an online social networking service that enables the users to send and read short 140-character messages called "tweets".

In the research, the author is going to use multilingual support of English and Finish. According to a report from the European Commission, most widely spoken foreign language of Europeans is English (38%). (European Commission, 2012) Additionally, to make the result of the research reliable, supporting the local language is a necessity.

Because, the main purpose of any business is to satisfy customers' needs, the companies need to deliver a good feeling to their customers. So, the author is going to use tweets and retweets volume in the research to monitor users' activity. Additionally, the author will analyse the user's opinions regarding the sample companies. Sentiment analysis, also called opinion mining, is not a new field of the research. The idea of sentiment analysis is as old as the idea of marketing. But it remains to be a popular and contradictory subject.

It seems that there are a lot of potential for a business in the social media analysis field. That helps them to know customers better, improve a marketing strategy, actively position their brand and improve corporate reputation. For this reason, the social media analysis subject it is very interesting to the author.

The need for real-time business intelligence and growing popularity of social media offer an urgent demand for this type of a research. Generally, the social media research related to marketing activities, sales, promotions, advertising. The author is diversely interesting in customers listen.

1.3 Research questions

The purpose of the research, as expressed in the previous section is to show the benefit of social media data mining utilization by companies to achieve better business results. To assist the purpose of the research, the next questions will be used.

The main research question: How may social media be utilized in strong reputation construction?

Sub questions about reputation building: How may online opinion data mining be utilized in reputation building? How do online customers' opinion impact reputation and business performance? Why is it needed to monitor and measure opinions in the first place? Is there correlation between online opinions and reputation?

Frame question about ethics of the social media analysis: What regulations, policies and ethical aspects regulate social media analysis?

Support questions: How does sentiment analysis impact the decision making outcomes? How valuable are these insights for business intelligence? How can a company use social media data for performance measurement?

The author presented the below mentioned questions into a graphical chart, "Coherence of research questions". That helps to see the logical dependence of information that requires an answer in the questions in sequence order.
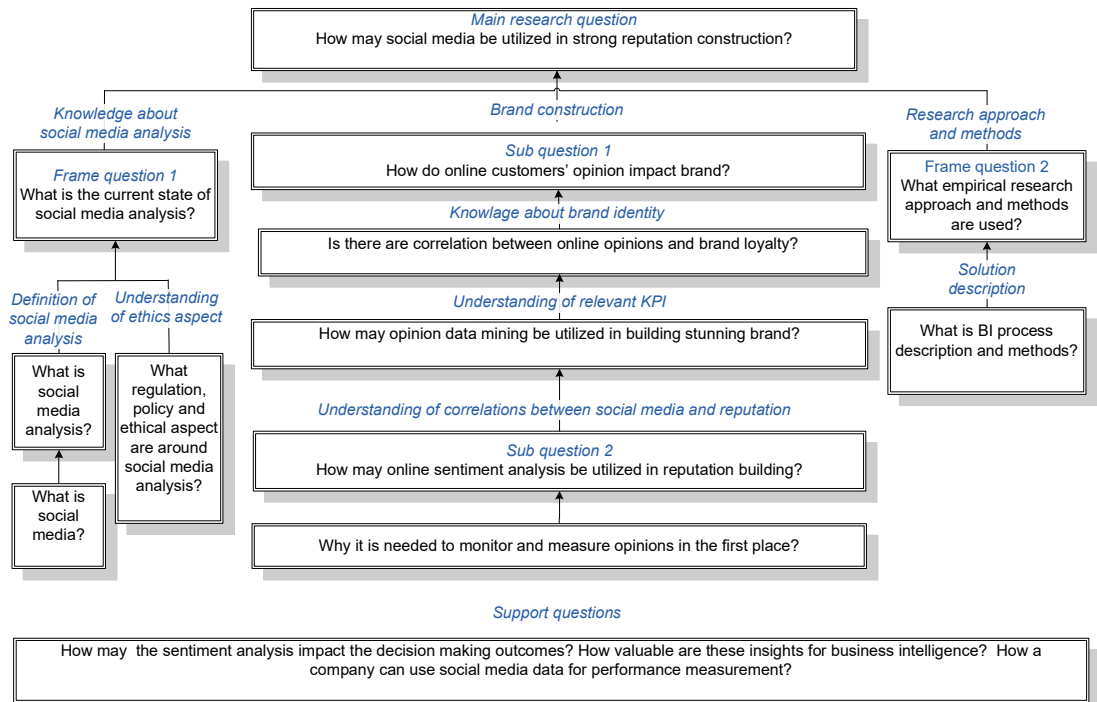
Figure 1: Coherence of research questions.

1.4   Limitation

> *All model are wrong, some are useful. (George Box, famous statistician)*

The main reason for the limitations of the research leans in the technology. A lexicon-based sentiment analysis is emotional classification of a text that done by machine, by use of keywords. The keywords classify the tweet post as "positive", "negative" or "neutral". The first limitation is a linguistic challenge of the way in which context may effect a meaning. Additionally, sentiment can be expressed without any direct use of the negative keywords. Other factor that may put the Lexicon-based sentiment analysis unreliable is that social media post can be expressed with sarcasm or irony. Otherwise, the sentiment analysis is not straightforward like any data analysis for that matter. Without proper analytical minds, it opens to substantial misinterpretations and subsequent poor decision-making.

1.5   Ethics of the social media analysis

The regulation of new technologies is generally a slow process. In nowadays, we are lack of legal guidance specific to a social media analysis from Finnish governments or EU directives. Not European countries or organizations, like CASRO or ESOMAR have different protections of human subjects and the preservering of data collection.

To ensure that the research meets ethical standards, the author will follow Twitter developer Agreement/Policy for social media analysis that covers human subject issues.

Integrity of Twitter product

- @username must be alwaus displayed with tweet text
- Do not modify, translate or delete a portion of the content.
- Do not aggregate, cache, or store location data and other geographic information.

Respect Users' Control and Privacy

- Do not store or publish non-public content such as direct messages
- Not allow to republish content by means other than via the Twitter API.
- Do not use a user's content to promote a commercial product or service. (Twitter, 2016)

1.6   Structure of the Thesis

The outcomes of the research are based on the data mining results of the twitter data. The literature will be covering books, articles, electronic sources and independent research results from themes of the business strategy management and technical aspect of the data mining implementation.

The four chapters of this thesis describe the different aspects of the research. Introduction, chapter 1, introduces the context, questions and

objectives. The methodology related literature review presented in chapter 2. Chapter 3 describes the data sources, methods of data collection and analysis. This chapter show the results of the data analysis. Conclusion, chapter 4, concludes this thesis and highlights the findings.

2 THEORETICAL FREAMWORK

> *The greatest obstacle to discovery is not ignorance-it is the illusion of knowledge. (Daniel J Boorstin)*

In today's digital world, many brands and organisations understand importance of utilizing the new technologies and available data. The question is how to approach the challenges and opportunities of available big data to utilize it in the customer-centric way to understand human behavior better. The nowdays technology makes things possible and available to everyone and peaple need to take advantages of it.

2.1 Big data

> *In God we trust; all others must bring data. (W. Edwards Deming)*

The Internet technology created big data term. A result of it is rapid growth of the structured and unstructured data. It is inferring to the large amount of data. Big data means the extremely huge data sets that have grown beyond the ability to manage and analyze them with the traditional data processing technique. In the late 2001 research report, made by META Group (now Gartner) the industry analyst Doug Laney defined big data as being three dimensional. Since then thouse reasonable characteristics been adopted globally: **V**olume, **V**ariety, **V**elocity. They quantify big data by taking into account the typical qualifiers around data nature. (Gil press, 2013)

Data *volume* represents the quantity of transactions, events, or amount of history that creates data volume. Typically, to analyze data, it is need to take a smaller data set, so called samples, to create data models. Additional, the data volume need to be classified into dimensions / attributes and measures, also called variables.

Data *variety* indicates to the twisty assortment of the data types. Traditionally, the operational data is structured. It is recorded into standardized SQL database based (Structured Query Language). Over

the past years, data has increasingly resembled "unstructured" as the sources of data become pictures, text, video, speech and image.

Data *velocity* illustrates that data is created in a real time. So it is speed characteristic that put demands on the businesses to process information with the near real-time response. (Minelli, Chambers & Dhiraj, 2013)

In 2012, Mike Gualtieri explains that three Vs are just measures of data — how much, how fast, and how diverse. But there are missing an estimated value definition of big data.

So, high volume, high variety, and high velocity are the essential dimensions of big data. But when you utilize big data to discover insights and predict future outcomes other characteristics of big data are equally important. Inderpal Bhandar, Chief Data Officer at Express Scripts noted in his presentation at big data Innovation Summit in Boston in 2013 that there are additional Vs. This second set of Vs characteristics includes *Value* and *Veracity*. They describe how accurate and useful is that data in creating advantageable value.

If we are looking in the internet for a definition of big data, we will find more interpretation. But they are mostly attached to V dimensions' interpretation. Because the data does not automatically mean value for business. There is notable difference between the data, information and knowledge. Considerable question is how to turn the data into their most valuable asset that is wisdom.

2.2 Social media

> *Social media spark a revelation that we, the people, have a voice, and through the democratization of content and ideas we can once again unite around common passions, inspire movements, and ignite change. (Brian Solis)*

The social media platforms, as a part of the internet and big data produce huge quantities of the data. This technology innovation is shaping the customer experience and offers the connection channel to them. Social

media are as old as human speech. But the web 2.0 technology solution has been changes the way how people communicate and exchange information. For sure, the social network is one of the greatest innovations of the past decade. Now, everyone can connect with a near- or long-distance friend, share photos and videos, tell the news, share ideas and resolve issues. Social network becomes the first point of contact.

Additionally, LinkedIn, Facebook, Twitter, and other social networks have an actual registration data provided by their users. This kind of information is the gold standard for the business. With this data, the companies can target their advertisements precisely based on job title, industry, company size, geolocation, and interests. The advertisements can also be targeted based on the user behavior. The power of display advertising is the ability to optimize, personalized advertisements and possibility to do it dynamically. But the next step is true understanding of the customer 's buying journey and the ability to model exactly where revenue comes from. (Russell & Sean, 2014)

According to Statista, the power of social networking is such that, the number of worldwide users is expected to reach some 2.5 billion by 2018 from 1.96 billion at 2015. That is around a third of Earth's entire population.

Another unique characteristic of social media is extremely high transparency that leads to immediate visibility of the consumers and the companies' activities. (Hennig-Thurau et al., 2013) On top of that, all the content can be passed through the social network very rapidly and efficiently. (Kohli et al., 2015).

Social media is the power channel of the customer's voice, that can't be ignored. In 2010, a student found clear trash bags with unsold clothes behind an H&M store in Manhattan. The student noticed that the clothes were damaged with a knife with intent to prevent from use anymore. She called to H&M with intention to get the explanation for their actions. But they refused to answer to the question. Why do they destroy clothes

instead to give them to charity? Next, she contacted to the New York Times. But H&M didn't bother to provide any comment for the article. However, his Twitter became the top 3 on the list of trending topics for a few days. After this, H&M representatives addressed the issue. They responded: "H&M is committed to taking responsibility for how our operations affect both people and the environment. Our policy is to donate any damaged usable garments to charity. We're currently investigating an incident in a NY store that is not representative of our policy." (Keane, 2010)

Social media is not only powerful; it is high-speed communication channel as well. In August 2015, Veet Finland published an unsuccessful 'Me Time' content marketing campaign on Facebook. In the post, there was a photo of a young woman in a bathroom eating ice cream and a quote 'Amusing night with three of my best friends: Ben, Jerry and Veet'.
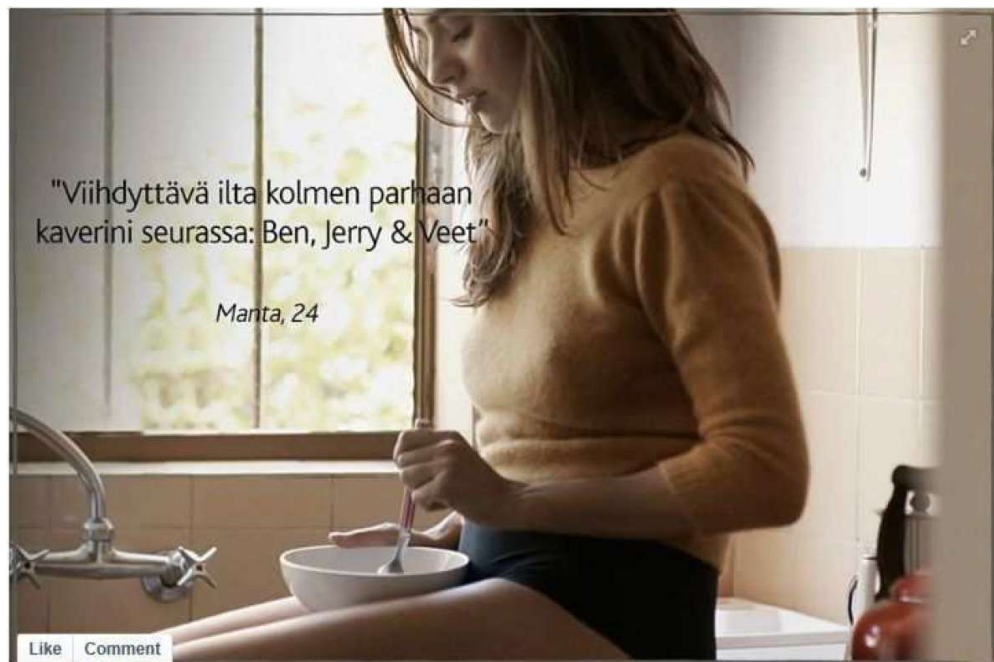

Figure 2: Veet's marketing campaign on Facebook.

The reaction of people was definitely unexpected by Veet Finland. The users immediately generated an avalanche of the post with sarcasm and anger.

Figure 3: Facebook's users respond on Veet marketing campaign.

Koho (2015) writes that the issue was that Veet wanted to create positive associations of enjoyable free time and waxing. Conflict was because the most people have experienced that waxing not fun or relaxing. Considering this, people may think their feelings are ignored. Additionally, after realizing the mistake Veet had made, they decided to keep pace with users' and react quickly. They immediately deleted all the content, including the users' comments. So, the result was, Veet asked to tell how users' nights go. And when people told, Veet did not desire to hear the

answers. Then they were in a situation where Veet is remembered, but not necessarily in the way they wished.

So, social media data may create huge business outcomes if a company utilizes this data in a right way. However, there are a number of challenges to successfully exploit these data possibilities. Managing and extracting value from social media data is a hard challenge that the companies may face because of a misinterpretation of it. This question is not only about technology they use, it is mostly about how they use it.

Kramer writes in HBR about this nowadays trend, that companies abuse social networks to share their news and information. He highlights that "the dominant form of communication is one-way –think tweets or status updates". The companies outsource the conversation with customers to Twitter or Facebook by doing the monologues conversation. Does this action meet company's objectives? Is it the right thing for the customer? The key of a successful dialogue with the customers is to start listening and tune into what the people are saying. He writes that "very few companies are able to use existing social media as a way to listen to customers". It can't be dispute the importance to shift from monologue to dialogue with customers. (Kramer, 2010)

Companies spend serious money by engage the people who maintain one-way monologue in social media although no customers participate in it. According to a Statista forecast, worldwide spend on online advertising gain from 25.14 billion U.S. in 2015 to 41 billion U.S. dollars in 2017. The study believes that the advertisers will continue to put more and more cash on the digital advertising over the next few years.

Therefore, the companies need to ensure that they have the systems that can "understand" customers and have a conversation accordingly. From the early days of advertising, marketers were trying to solve the fundamental advertising question: is this advertisement reaching the right audience?

Russell and Sean write that nothing new to concentrate focus on customers and use data to better understand them. They underline that every company has access to a goldmine and prospect data of the valuable customer in nowadays. The key is to find out what is important in this data so the companies can take action before any of their rival do. The companies that most effectively use these approaches to gain insight into their customers and act on that data will win. (Russell & Sean, 2014)

While the advertisement publishing still plays a key role in terms of brand awareness and new product discovery. Getting the most out of social media requires a greater investment in deeper research to find information that is truly relevant to customers needs. (Russell & Sean, 2014)

To drive results with the right tactics and appropriate planning company needs to build a social media strategy that works. It is all about listening and targeting the right people with the right messages. For this understand of the audience is compulsory.

2.3   Data value creation

The natural consequence of the Internet and web 2.0 new technologies is data explosion. That dramatically changed the way of companies' function. The leverage properly becomes a data management. In Fact, data alone cannot help decision makers to make a decision. Instead, too much data or bad quality data can often mislead and distract the decision-making process. So, raw data collection needs to transfer into action-guiding wisdom.

Russell Ackoff, a system thinking theorist and professor of organizational change, classified human's mind into five categories. They are:

1. *Data* is a product of observation and consist of symbol, events and properties.
2. *Information* is the data that are processed to be useful. It is providing answers to the questions "who", "what", "where", and

"when" and "how many". The data is only valuable as the information.

3. *Knowledge* is awareness – "I know you are there" - and ability- "I know how to drive a car". Intelligence is the ability to acquire knowledge.

4. *Understanding* is appreciation of "why" questions. Understanding is required the relevance of data and information for analyse. It is need to understand why the situation is what it is and how its characteristics are causally related to our objectives.

5. *Wisdom* is evaluated understanding. "Wisdom is the ability to perceive and evaluate the long-run consequences of behavior."

So wisdom is an outcome of understanding, which is an outcome of knowledge, which is an outcome of information, which is an outcome of data. The first four categories of data related to the past. They are information, knowledge and understanding. The categories deal with what has been happened or what we already known. Only wisdom deals with the future because it incorporates vision and design. Only by wisdom, people can have understood of future rather than just see the present and past. For achieving wisdom, the people need successively go through all previous categories. (Ackoff, 1999)

Structural representation of relationships between data, information, knowledge, and wisdom has created several phenomena identifications.

Mohanty, Jagadeesh and Srivatsa describe the data transformation model from raw data into action-guiding wisdom as a pyramid. (Mohanty, Jagadeesh & Srivatsa, 2013)
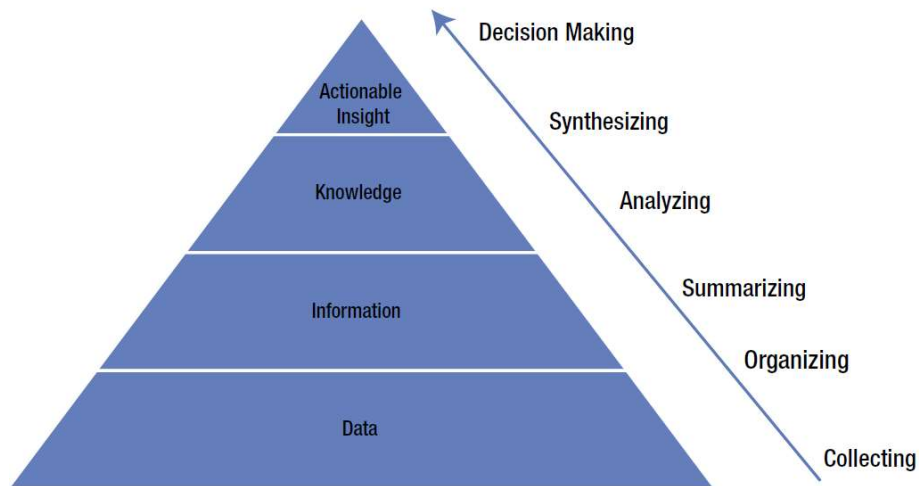
Figure 4: Transforming raw data into action-guiding wisdom. (Mohanty, Jagadeesh & Srivatsa, 2013)

The wisdom pyramid, also known variously as the wisdom hierarchy, the knowledge hierarchy, the information Hierarchy and the knowledge pyramid.

2.3.1 Data-driven decision managment

> *Data-driven decision management (DDDM) is an approach to business governance that values decisions that can be backed up with verifiable data. (Data and data management glossary, 2016)*

Innowdays, market driven information technologies produce more digital data that has ever available before. So, companies in almost every industry are focused on exploiting this data for competitive advantage. According to McElheran and Brynjolfsson, those phenomenon leads that many managers change the way how they make decisions. They start to practice of basing decisions on the analysis of data rather than purely on intuition. (McElheran & Brynjolfsson, 2016)

There is strong evidence that business performance can be improved substantially via a data-driven decision making. The detailed survey of 179 large publicly traded firms, performed by Brynjolfsson, Hitt and Kim, shows

"that companies that adopt DDD have output and productivity that is 5-6% higher than what would be expected given their other investments and information technology usage". (Brynjolfsson, Hitt & Kim 2011)

According to McKinsey research, big data has a potential to add value to a business across all industries. Big data are helping businesses to become smarter, more productive, and better at making predictions. (Manyika, J., et al. 2011)

Mohanty, Jagadeesh and Srivatsa wrote in 2013, that especially next business functions that can be easily enhanced by a big data program are:

- Finacial serices
- Supply chain, logistics and manufacturing:
- Online services and web analytics
- Energy and utilities
- Madia and telecommunication
- Health care and life sciences
- Retail and consumer product

	Volume of Data	Velocity of Data	Variety of Data	Under-Utilized Data ('Dark Data')	Big Data Value Potential
Banking and Securities	High	High	Low	Medium	High
Communications & Media Services	High	High	High	Medium	High
Education	Very Low	Very Low	Very Low	High	Medium
Government	High	Medium	High	High	High
Healthcare Providers	Medium	High	Medium	Medium	High
Insurance	Medium	Medium	Medium	Medium	Medium
Manufacturing	High	High	High	High	High
Chemicals & Natural Resources	High	High	High	High	Medium
Retail	High	High	High	Low	High
Transportation	Medium	Medium	Medium	High	Medium
Utilities	Medium	Medium	Medium	Medium	Medium

Figure 5: Big data value across industries. (Mohanty, Jagadeesh & Srivatsa, 2013)

Because data and a value for a business they are two different things, companies need to extracting useful knowledge from data. Business success in today's data-oriented business environment requires being able to think about how available data could be apply to the particular business problems. The biggest challenge and obstacle is to implement this consept into practice.

Practical approach to a data manipulation concept represents Han, Pei and Kamber. It is a prospect for companies to become capable of processing data in a flexible way. The process of knowledge extraction

from the data also called as a data mining. The term of "data mining" emphasize mining of the prosses importance. Data mining is alsow used to say knowledge discovery from data, or shortly KDD. Practical approach of data mining is an essential step in a process of knowledge discovery. Knowledge discovery as the process represented in Figure2 and consists of the following steps:

1. Data integration and cleaning, alow called ETL from Extract Transorm and Load data.
2. Relevant data selection for the analysis task.
3. Data transformation or consolidation into forms appropriate for mining by performing summary or aggregation operations.
4. Data mining is a vital process where business intelligence methods are applied in order to extract the data patterns.
5. The patterns presentetion and evaluation.
6. Visualization and knowledge representation of mined knowledge to the user. (Han, Pei & Kamber, 2006)
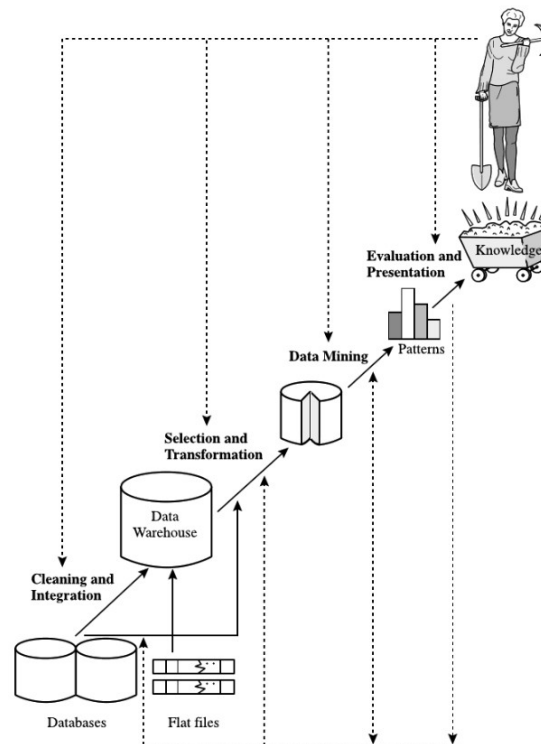
Figure 6: Data mining as a step in the process of knowledge discovery. (Han, Pei & Kamber, 2006)

2.3.2   Social media analysis

> *To be able to improve something, you have to know how to measure it. (Greg Brue)*

In facts, the social media industry is undergoing growth. We have millions of devices, systems and users connected on the Internet generating these kinds of data. Social media is not only a critical communication tool with customers for any brand across all industries. The demand for real-time business analysis and growth popularity of social media offers the great opportunities. Social media data help get real-time insights on business performance, customer experience and company reputation. People are talking about your company or brand online, whether you choose to listen or not. But managing social media data and understand the implication is a challenge advantage for the companies. They may struggle to utilize social media data to get maximum potential.

According to Blanchard, one of the most common mistakes made by the companies when they are planning to take social media in use is that they focus too much on the social media tools and the platforms and not enough on their business objectives. (Blanchard, 2011)

In 2011, Lovett writes that measuring social media is a challenging task and rewarding at the same time. The main challenge is to identify the goals and metrics that have meaning and matter to the business because measurement itself is not the goal. Other is collaborating with the stakeholders to align the metrics that will truly identify progress toward their business goals. Next one is finding the right tools and channels. Take the time to identify channels where your audiences are because most popular chanel may not necessary work for every company. (Lovett, 2011)

Blanchard in 2011, found that there are the top five business functions that can be easily enhanced by a social media program: sales, customer support, human resources, and public relations. (Blanchard, 2011)

Lovett in his work defined the outcome metrics and the business value metrics to aid the social media program. For social media network they are users, age, gender, location, likes, tweet, and retweet, comments, fans, post, and sentiment and so on. (Lovett, 2011)

There are some issues accosiated with cosial media as well. According to Nexgate research, one of the biggest social media security issue is spaming that rised only at half of 2013 in 355 %. Social media spam is irrelevant or unsolicited messages for purpose of advertising, phishing, spreading malware, etc. The cosial media spam "can damage brand appearance and turn fans and followers into foes". (Nexgate, 2013)

The one way of capturing social media data and making better sense of it is sentiment analysis. Recognition of the emotional value of social media enables the business to gain deeper insight through the analysis of unstructured data. Because sentiment analysis is a complex and challenges task even for humans, this type of analysis discovered deeper in next chapter.

2.3.3 How sentiment analysis works

Sentiment analysis refers to a data mining and test analytic techniques that determine an emotion behind a social media text. Is is also called opinion mining or emotion analysis. (Kaur & Gupta, 2013) Sentiment analysis detects the subjectivity and polarity of the text. It is classifying the social media text as positive or negative opinion. (Taboada et al., 2011).

Lui (2012) writes that opinions are key influencers of human behavieors. Individual customer wants to know the opinions of existing users about their brand, product and services before purchase decision. Because of this, business and organisations always need to find customers or public opinion.

Several studies have shown that the users opinions in social media may used as a predictor of consumer choices indicator. E.g. Mishne & Glance (2006) demonstrated that positive sentiment in social media correlate with the product's financial performance by proving the good interdependence between the positive sentiments and the movies' financial success.

There are different approaches available to polarity classification. They could be distinguishing between supervised machine learning and unsupervised lexicon- based approaches. (Fernández-Gavilanes et al., 2016) The unsupervised lexicon- based approach is a mostly used way to perform a sentiment analysis. According to Taboada, lexicon-based sentiment analysis is a robust and result is good. (Taboada et al., 2011).

A sentiment lexicon, also called a polarity or opinion lexicon consists of a list of words with associated values representing their sentiments. These values are usually integers expressing polarity and polarity intensity as increasing or decreasing absolute values. (Hu & Liu, 2004)

Several limitations are associated with lexicon based sentimental analysis that challenges the accuracy of the result.

- An interrogative or a conditional sentence containing sentiment words may not express any sentiment. Like, "If I can find something very good looking dress, I will by it." (Lui, 2012)
- A positive and negative word in the sentence may have opposite orientation. A simple example is when you ask someone young person "How a you?" He may respond "I'm sick" that mean he are fantastic. (Lui, 2012)
- Social media post may not contain any sentiment words or express with sarcasm or irony. (Lui, 2012)

Sentiment analysis is chalanges task, at the same time it is huge opportunity for business. It may help companies to evaluate the success of campaigns in real-time and demand for future action. By doing so, tremendous market research avoided. It helps to evaluate the sentiment

developments over time about a brand or a product and benchmark them against competitors. (SSIX 2016)

2.4   Strategic Business Management

It is may be difficult to apply to the big data puzzle without appropriate strategy. Johnson and Scholes defined strategy as a term of concept of development direction: "Strategy is the direction and scope of an organisation over the long-term: which achieves advantage for the organisation through its configuration of resources within a challenging environment, to meet the needs of markets and to fulfil stakeholder expectations". (Johnson & Scholes, 1993)

The business world changes all the time during continuous evolution, especially last years. That demands new innovative and competitive business strategies that attract customers and make customer loyal; to make sure that consumers choose your product or services over others. This task becomes more and more challenges because of increasing customer expectations, large competitions in the market and volatile economy. For this, companies need to create a pipeline of innovative and competitive advantages for customers to replace those that have been competed away.

As a result, the worlds of business strategy and innovation have gotten much closer to one another for addressing how a company will act in the face of uncertainty. To innovate and strategize company must ask new questions, and see in the whole new ways. Strategy is about long-term thinking. In addition, it need accommodate the uncertainties and volatilities inherit from the complex and dynamics live to control company destiny.

In 2007, Zook introduces the ways how to expand the basic business and facing challenging times. He offers five diagnostic questions to evaluate its core business. Thay are relative to customers, key sources of differentiation, profit pools, capabilities, and organizational culture of your core business. The main line is that the sustainable growth could be

achieved when company is able to combine unused capabilities with the existing capabilities for producing something new and better. By finding the hidden assets, that are neglected businesses, unexplored customer insights, latent capabilities. (Zook, 2007)

Sometimes it is hard to know when company core business must change. For this management teams can learn to recognize early signs of erosion. Ofcouse challenges can be simply a temporary failure in execution. Because of this Zook address three reasons for rethinking the core strategy. The first is profit pools, is company is targeting a shrinking or shifting profit pool. The second one is the nowadays trend reason is inherently inferior economics. The third reason is not maintainable business environment. Like company, who manufactures cell phones, might find that competitors replicate its once unique source of differentiation and target market is saturated. (Zook, 2007)

Importance of differentiation or being different in the marketplace from rivals was emphasized by Porter in 1996. He defined this in clever way, that competitive strategy produces unique value through differentiation by choosing different activities and monitor how well they are performed. By strategic positions, he means, performing different activities from rivals' or performing similar activities but in different ways. Performing similar activities better than rivals, he refers to operational effectiveness. Porter emphasize that operational effectiveness is not the strategy. So, the companies must distinguish between the operational effectiveness and the strategy because concentration on core competencies and competitive positioning via benchmarking can lead companies to the poor outcomes. (Porter, 1996)

In 1979, Porter introduce "The five competitive forces that shape strategy". He explains that the most successful businesses find the niche in the market where the force/s are the weakest.
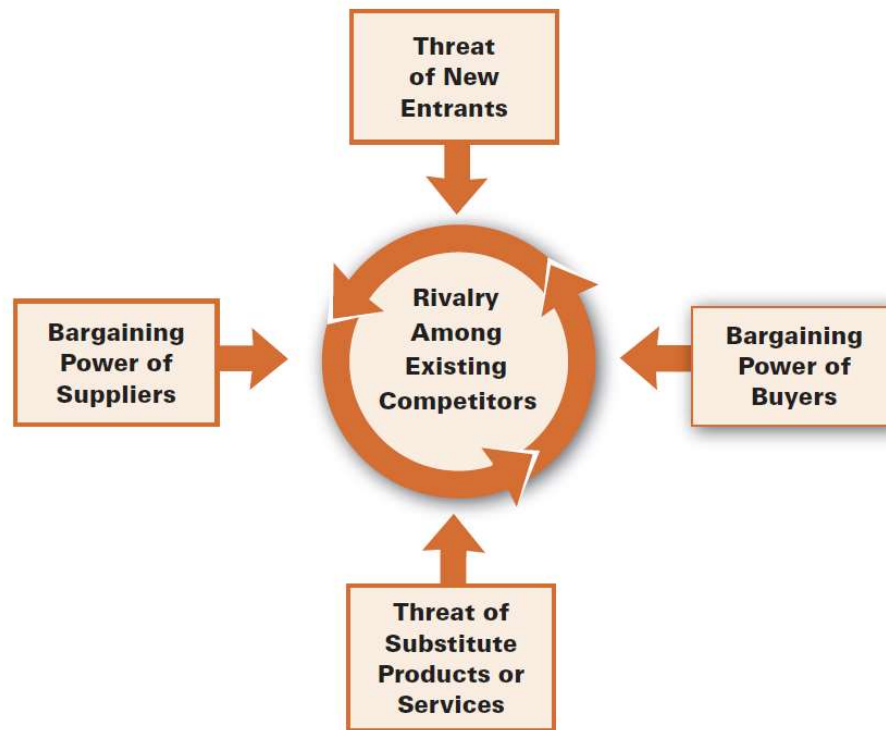
Figure 7: "The five competitive forces that shape strategy" (Porter 2008)

He finds that all industries have different profitability regarding return on invested capital. There are connection between combinations of forces that create a low or high return. The clear link is if these forces are too intense in an industry, it is very hard for a company to earn a return on investment. A low return industry does no investment. A low return industry does not demand that all five forces be powerful, but one force can disturb the profits in an entire industry. In other point, the five competitive forces are threat of entrants. The threat of entry depends on the height of entry barriers including power of suppliers, power of buyers, threat of substitute products or services, and rivalry among existing competitors. (Porter 2008)

Kim and Mauborgne (2005) divided the markets into the red and blue oceans. By the red oceans they means competing in the same market space, and by the blue oceans competing in a new, previously unknown market space. Correspondingly, industrial rules are different. Because in

blue oceans customers' demand is just created that create the opportunity for profitable and rapid growth. In red oceans, where industries already exist, companies compete by gobbling for a greater share of limited demand. As more rival market space is the more crowded, prospects for profits and growth decline. But the most exciting thing is that you can create the blue oceans by two ways. One is to launch completely new industries. Like eBay did with online auctions. But the most common way is to create the blue ocean from a red ocean by expanding a company's boundaries of an existing industry.

Red Ocean Versus Blue Ocean Strategy

The imperatives for red ocean and blue ocean strategies are starkly different.

Red ocean strategy	Blue ocean strategy
Compete in existing market space.	Create uncontested market space.
Beat the competition.	Make the competition irrelevant.
Exploit existing demand.	Create and capture new demand.
Make the value/cost trade-off.	Break the value/cost trade-off.
Align the whole system of a company's activities with its strategic choice of differentiation *or* low cost.	Align the whole system of a company's activities in pursuit of differentiation *and* low cost.

Figure 8: "Blue ocean strategy" (Kim & Mauborgne 2005)

The interesting study about the drivers of profitable and sustainable corporate's growth, was performed by Zook and Allen. The study found that the most successful companies were able to consistently and profitable outgrow their business by developing a formula for rising of their core businesses in the predictable and repeatable ways. They may continue to focus on their core products, but sell thea products/services in new geographical areas, through new distribution channels, or to new customer segments. Such companies develop and apply a strict repeatability formula to these adjacency moves. This formula enables them to change just one variable at a time and execute moves faster. The

successful repeaters in the study have two common characteristics. First, they were extraordinarily disciplined, applying rigorous screens before they made an adjacency move. And second, they developed their repeatable formulas by studying their customers and their customers' economics very, very carefully. (Zook 2009)

McGrath and MacMillan (1995) write about importance in approached innovative ventures with the right planning and control tools. They discover the importance of having conventional planning methods for the new venture. That move ventures through a series of milestones, rather than trying to plan them all at once. It is means that innovative strategy needs to be a systematic, ongoing process with a set of tools and processes that allow company achieves the desired results. Thay describe the differences between discovery-driven planning that turns the assumptions into knowledge along the process and platform-based planning that treats the assumptions as facts instead of best-guesses. (McGrath, 1995)

In 2010 McGrath writes, that strategies are not traditional skills of planning and rock-ribbed execution any more in nodays highly uncertain, complex and fast-moving environments. It is about about quick getting insight, rapid learning and experimentation. (McGrath, 2010)

2.5   Corporate reputation building

> *It takes 20 years to build a reputation and five minutes to ruin it. If you think about that, you'll do things differently. (Warren Buffett)*

There are a lot of different definitions of corporate reputation available. "Reflects the history of past actions." (Yoon et al.,1993) "A corporation's values." (Smythe et al., 1992) "A set of economic and non-economic attributes". (Weigelt & Camerer, 1988) Perhaps the biggest obstacle to the creation of one uniform of the concepts is mixing of items: identity, public image and reputation because they are closely related to each other. (Barnett, Jermier & Lafferty, 2006)

However, the research performed by Barnett, Jermier and Lafferty (2006), evaluate the sole definitional statements of the corporate reputation find the three distinct dimentions of reputation among the definitions available. There are a state of awareness, reputation as an assessment and reputation as an asset. Aula & Heinonen (2016) writes, reputation does not exist outside of time and represent always present situation. Reputation defines a company's past heritage, present activities results and expectation for the future.

Corporate reputation has become popular in last years because there are a lot of given evidences that show the correlation between the corporate reputation and various intangible and tangible benefits. (Graeme, Ronald & Cary, 2012) Fombrun in his book writes about the benefits of a good reputation. That reputation is a competitive advantage and the most valuable intangible asset a company holds. (Fombrun, 1996)

Studies of the Fortune 500 companies have shown that reputation is associated with a company's financial performance. Thus the companies with good reputation have about 12 percent higher price to earnings ratios. (Graeme, Ronald & Cary, 2012)

The good reputation enables to attract customers easily and manage loyality of the existing ones without extra financial investments. It helps to attract the best employees in the labour market and manage crises like bad economic times more smoothly. (Aula & Heinonen, 2016) Wartick (1991) write that the coverage of any form of media has an impact on the corporate reputation and consequently of the value of company from the investor's point of view.

Aula & Heinonen write about research results that use broadly collected data from 2001 to 2009 from Finnish companies that have similar reputation profile. The research classifies companies into "archetypes" by using given characteristics. (Aula & Heinonen, 2016)

- The **solid** reputation companies have stable and high level reputation, e.g. Kone. Kone is the global leader in manufacture of the elevator and escalator industry.
- The **varying** reputation companies' reputation varies without any clearer trend, e.g. Finnish shipping company Finnlines.
- The **deteriorating** reputation companies that reputation is on the decline, like Elcoteq that was a Finnish consumer electronics contract manufacturer, filed for bankruptcy on 2011. (Aula & Heinonen, 2016)

The research results show that only the five companies out of 100 were able to maintain a solid reputation and return on investment (ROI) of the companies with a solid reputation was much higher. (Aula & Heinonen, 2016)

Archetype	Return on investment (ROI)
Solid	29.95
Varying	12.61
Deteriorating	2.97

Figure 9: "The return on investment (ROI) in three different reputational archetypes, 2001-2008" (Aula & Heinonen, 2016)

The Internet increased significantly the number of available channels and a way of communication. Customers nowadays have more proactive utilization and more control over the communication process. They have a possibility to find information globally, contacting other customers, or expressing their own opinions in social networks. So, the corporate reputations are not about what companies do or say, it is about how the customers see and respond to company actions. (Bunting & Lipski, 2001)

Furthermore, the neuroscience studies confirmed that there is strong connection between emotions and cognitive processes exist. Emotions represent a complex reaction that influences human behavior. Studies show that humans cannot even think without emotion. Emotions influence

all of our decisions, as well purchase goods decisions. Because, there are a complex relationship exists between emotion, behavior and disease. (Beesley, 2005)

The digital reputation (DIRE) study conducted at the University of Helsinki and Aalto University in Finland between 2010 and 2012 affirms that unconscious emotional reactions to reputation directly affect the decision-making, cognitive and emotional processes. The reputation opinion is a consequence of evaluations and experiences made by individuals and public opinion. (Laaksonen et al. 2012)
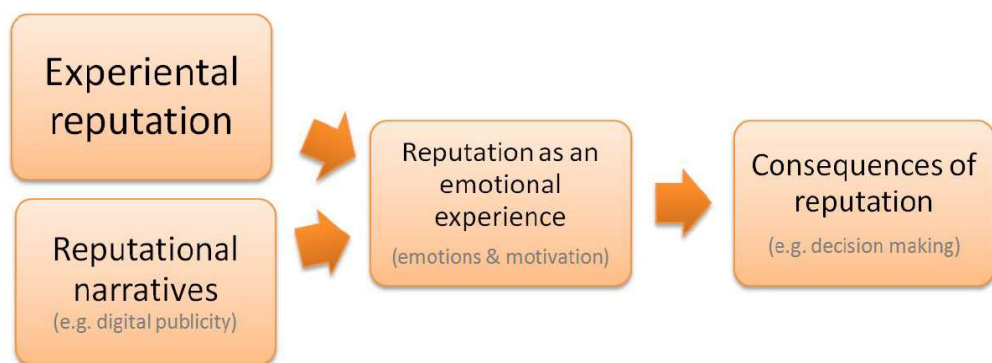

Figure 10: "Two paths of reputation formation." (Laaksonen et al. 2012)

The DIRE-study demonstrates that company-related content in social media has serious effects on a company's performance in the long run. The negative online publication is not only bad for all companies but it is extremely harmful if company has already a poor reputation. (Laaksonen et al. 2012) Corporate reputation may also be a critical factor in responding to a crisis. (Schnietz & Epstein, 2005).

Aula and Heinonen (2016) recommend companies to create the strategy to produce the desirable online discussions. The companies should not only take into high consideration the negative online discussion but also provide the stakeholders with emotionally positive and satisfying experience. "The reputation cannot be built but it needs to be earned."

3 RESEARCH APPROACH, METHODOLOGIES AND EMPIRICAL STUDY

Web 2.0, social media and social computing technologies provide not only information; additionally, they allow users to interact with these web pages. It allows anyone to create and share information about frustration, dream or wishing, collaborate and build relationships. Social media platforms are web applications that the users can use for this. The author of the research used Twitter platform to hear customers' voice and find the ways to make social media work for business.

In research, the author used a quantitative research method. Quantitative method is used in collected online data analysis and in analysis of literature review of existing research in the same subject's area.

3.1 Research approach to data source

Twitter posts, called tweets is 140 characters short and authors usually are straight to the point. Because of this, the posts are easier to analyze and achieve better sentiment analysis accuracy. (Lui, 2012)

Twitter version 1.1 provides two APIs that are REST and Streaming. Streaming API requires keeping a continuous HTTP connection open and it provides data in almost real-time. The REST APIs supports the short-lived connections, they are rate-limited and does not make the data available older than a week. The REST API has a language parameter that can be set to a language code, eg. 'en' to collect English data. (Twitter 2016)
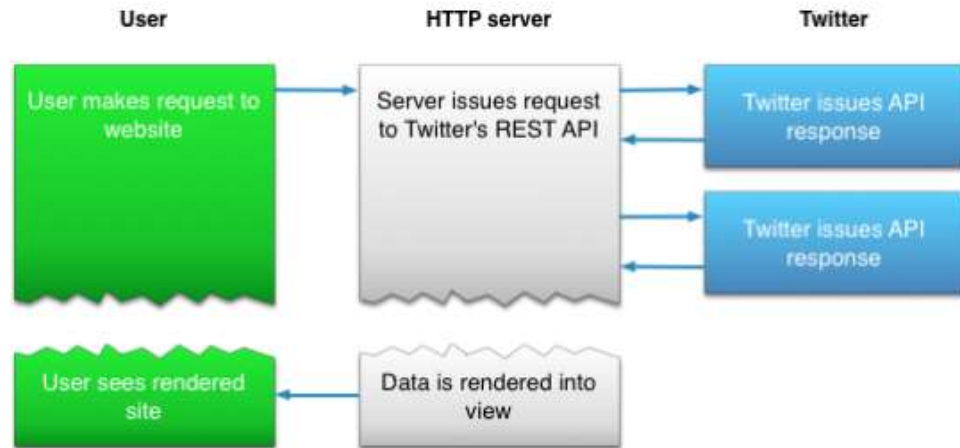
Figure 11: Twitter's REST API. (Twitter 2016)

A REST API of Twitter was used during data collection. For this, the interface to REST connection API of twitter was created by using the Ruby programming language. The author extracted the twitter social media stream data through OAuth authentication for future data collection and mining.

3.1.1 Data collection

This section gives an overview on the data collected that served to answer the research questions. Because objective of thesis is to analyze the Twitter messages posted in purpose to show benefit of utilize social media data mining in company reputation managment. Only tweets accociated to selected companies for period 24.2.2016-1.8.2016 are collected by using search words presented in Appendix 1. Data size of the collected data is 157103 rows or tweets.

3.1.2 Companies' presentation

Data collected from 20 companies during period about half a year. Only the companies with notable findings evidence are presented in this chapter. Those companies' data are used in the research finding chapter. The few companies added to data analysis after entire data collection was

started. The author added Nordea to data collection on 9.4.2016. It started to be very interesting because of the challenges with resolving "Panama Papers" leaks. List of all companies can be found in Appendix 2.

Tallink Silja Line is the leading provider of leisure and business travel and sea transportation services in the Baltic Sea region with over 50 years of operating and cruising experience. They operate with 16 vessels and five hotels. Their revenue was 945 miljon euros in 2015. The company employs nearly 7 thousand employees. They serve 9 million passengers and transport 308 thousand cargo units annually. Tallink Silja Line is listed on Nasdaq OMX Baltic, stock symbol TAL1T. The company is targeting to be the leading provider of mini-cruise, passenger transport services and leading provider of ro-ro cargo services in all the regions the Group operates, as well as extending operation wider to other coutries. They aimed to be the market pioneer in Europe by offering excellence in leisure and business travel and sea transportation services. (Tallink Silja Line, 2016)

Viking Line is a market-leading brand in passenger traffic on the northern Baltic Sea offering both recreation and cargo carrier services. They operate with 7 vessels and collaborate with quality hotel options in several price categories. The revenue was 530.5 million euros in 2015. During 2015, the average number of employees in the Viking Line Group was 2,735. They serve 6.7 million passengers and transporting 100 thousand cargo units annually. The company is listed on Helsinki Stock exchange with the stock symbol VIK1V.HE. Viking Line's strategic objective is to be the market's best value for money by providing good quality at affordable prices. The company target to offer selective quality factors are friendly service, fully functional and clean facilities, good food, enjoyable entertainment and attractive shopping. They aim at large travel volume and high capacity utilization. The sources of income are ticket, cargo and on-board sales. (Viking Line, 2016)

Nokian Tyres is a tyre manufacturer. They focus on products passenger car, trucks, and heavy machinery tyres for demanding conditions and

customer requirements. Nokian Tyres has been the leading brand of winter tyres for more than 80 years. They operate in Finland, Russia, Central Europe and North America. The net sales were 1,360 million euros in 2015. At the end of year 2015, number of employees in the Nokian Tyres was 4,400. Their strategic factors are demand conditions, innovateive core competence, customer satisfaction and market. The company is listed on Helsinki Stock exchange in 1995 with the stock symbol NRE1V.HE. (Nokian Tyres, 2016)

Wärtsilä is a Finnish corporation which manufactures and services power sources and other equipment in the marine and energy markets. The core products of Wärtsilä include large combustion engines used in cruise ships and ferries. In 2015, the company employed 18 857 workers in more than 70 countries with revenue 218 miljon euros and it is headquartered in Helsinki. Wärtsilä operates globally but its Ship Power division is heavily focused on Asia. The company's strategic objectives are support sustainability performance through the continuous improvement process and manage operating risks. The company is listed on Helsinki Stock exchange with the stock symbol WRT1V.HE. (Wärtsilä, 2016)

Nordea is the result of the mergers of the Finnish, Danish, Norwegian and Swedish banks that took place between 1997 and 2000. Nordea's main operational centres are in Copenhagen, Helsinki, Oslo and Stockholm, and with regional offices also in Estonia, Latvia, Lithuania, Poland, Russia, Singapore and USA and headquartered located in Stockholm. During 2015, the average number of employees in the Nordea Group was 29,815. The total operating income was 10,140 million euros in 2015. Value captured by addressing three key strategic focus area advisory, digital experience, efficiency and scale. The company is listed on Helsinki Stock exchange with the stock symbol NDA1V.HE. (Nordea 2016)

Metso Corporation is a Finnish process industry sector that is listed in Helsinki Stock exchange with the stock symbol MEO1V.HE. The company operates in three business area minerals capital, minerals service and flow control. Its customers include mining and stone crushing industries and oil

and gas companies operating in the fields. Metso was born in 1999, Valmet Corporation and Rauma Corporation's merger into a single company, and employs approximately 16 000 people in over 50 countries. Net sale was 2,923 million euroa in 2015. (Metso 2016)

3.1.3 BI process description

The complite BI tool has been build for proper data analysis purpose. That extract selected data from Twitter. During the next phase, called ETL, extract data cleaned and classified. Transformed data loaded into MySQL database for retention. Twitter's collected data from MYSQL database visualized by using Qlik Sense data discovery tool.

The graphical representation of the whole BI process shows into Figure 12.
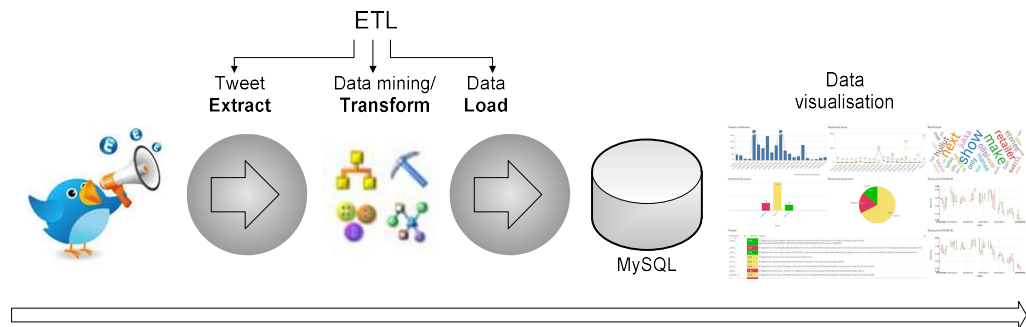

Figure 12: BI process.

Data for the research extracted from Twitter by using search words presented in Appendix 1. After extraction, data cleaning and mining performed during ETL process. The data mining consists of sentiment classification of tweets.

For this purpose, an unsupervised lexicon-based method was used for sentiment analysis in English and Finish languagies based on sentiment lexicon words lists.
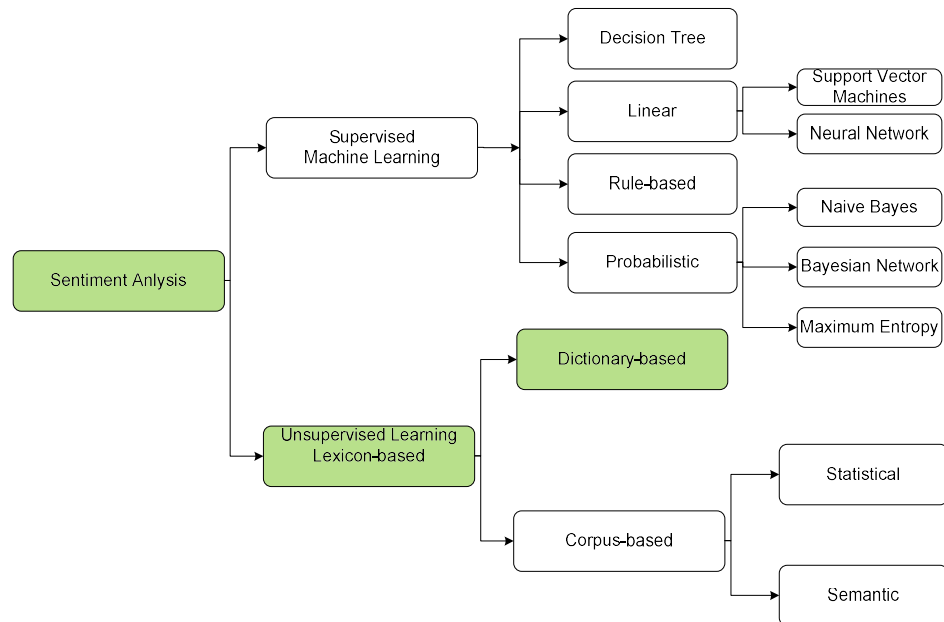
Figure 13: Sentimental classification techniques.

The sentiment lexicon lists of the words presented in Appendix 3 and Appendix 4. They take into consideration the linguistic content of messages. Additionally, the list of emoticons with emojis was used that can be finding in Appendix 5. They were used in text parsing stage to determine the polarity of tweets.

Twitter supports over 1000 emojis, such as smileys, love icons, animals and so on. The emoji is a form of language that are not used in the formal text language. The emoji intended to help in expressing and optimizing thoughts and feelings in electronic communication. As an example, emotional context of the neutral, simple phrase "I'm going home" easily changed by using facial emoji.

Text	Emotion	Note
I'm going home	:-)	Happy smile
I'm going home	:-(	Sad frown
I'm going home	:-||	Angry
I'm going home	;-)	Wink
I'm going home	:-D	Deep laugh

Figure 14: Emoji usage

To raise the accuracy of the research, spam messages were removed from collected data into two stages. Brand-owned social media account

tweet messages were removed during data mining stage. They were identified by comparing account name to search words. Those tweets were removed where they were the same.

if tweettext.downcase.include? search_word.downcase

The rest of brand-owned social media account tweet messages were removed before data visualization, during data load into application. It is done because they mostly include advetaising and promotin message. The author used customer's voice as a source of information in the research.

BI application has been built for a company's data visualization by using Qlik Sense data discovery tool that presented in figure 15. The BI application includes six gauge charts for assisting in the company data analyse. Color labeling are used for presentation of twitter posts' sentiment classification. Red color represents negative comments, green color shows positive and yellow color is neutral. Additionally, stock price embedded extension is used to assist Twitter's data analysis.

Figure 15: BI application.

The next chapter includes furthermore information regarding data analysis methods.

3.1.4   Analysis of the collected data

The author used the next key-performance indicators to gain insight in the volume of social media messages regarding a company and leverage social media utilization benefit.

- Total number of tweet mentions about a brand that shows how often is your company mentioned on twitter
- Total number of retweet of conversation about a brand
- Total number of positive conversations or mentions about a brand e.g. a tweet saying: I love brand ABC
- Total number of negative conversations or mentions about a brand e.g. a tweet saying I hate brand ABC
- Negative ratio or percentage of negative conversations / total conversation

$$P = \sum_{T1}^{T2} N / \sum_{T1}^{T2} X$$

    Where P is negative ratio, N is negative tweets, X all tweets, $T1$ start date, $T2$ end date.

- Positive ratio or percentage of positive conversations / total conversations

$$D = \sum_{T1}^{T2} P / \sum_{T1}^{T2} X$$

    Where D is positive ratio, N is positive tweets, X all tweets, $T1$ start date, $T2$ end date.

Additionally, "word cloud" has been used to help in analyzing the most frequently used terms in the tweets. Word cloud is a way to visually represent the content of a block of the tweets. The word cloud picks out the most common words found in the tweets. It is shows the words in a way that lets the viewer know what the words are used in a block of tweets and with what kind of frequency.

Stock market curve retrieves stock quotes from Yahoo! Finance: http://finance.yahoo.com/. It was used to assist data analysis. The Stock

market curve shows the stock price change of the particular company over the twitter's data collection time.

To support finding results, quantitative method was used in analysis of literature review and of existing social media research in the same subjects' area.

3.2 Research results

*Human behavior flows from three main sources: desire, emotion and knowladge. (Plato)*

This chapter presents the most important findings from the research social media analysis. The findings and the analysis are presented in written form. Additionally, printscreen from the created application is also used. The general reputation status of the cases companies is determined by Twitter's data analysis. The data analysis follows offline and online news about the case company.

3.2.1 Silja Line vs. Viking Line

Numbers by themselves do not tell a lot, we always need to compare them to see the difference. So, the first case presented social media analysis from two companies Tallink Silja Line and Viking Line. The purpose of this case is to see difference in social media activity and compare it to stock price change. The companies are competing in the same cruise ferry operation industry, where there are only a few service providers in the northern Baltic Sea region. More information about companies may be found in chapter 3.1 2 Companies' presentation. Service offering between companies seems to be really similar and this is reflected in similar price levels. The author made a strategy canvas to visualize the company's strategic profile or value curve. That was based on personal experience assumptions and annual reports of the companies. The author has the experience that companies' price more or less the same, quite standard. Also the destinations, frequency of departures and travel durations are similar. Both, Viking Line and Tallink Silja Line have their own loyalty

programs which are quite equivalent. Tallink Silja Line give consumers benefit by getting S-Group bonus points and Viking Line offer discount from cruises.

Strategy canvas demonstrates how companies identify competition factors they strategically invest in. (Chan & Mauborgne, 2002) The strategy canvas clearly shows the strong competition in the industry between Viking Line and Silja Line.
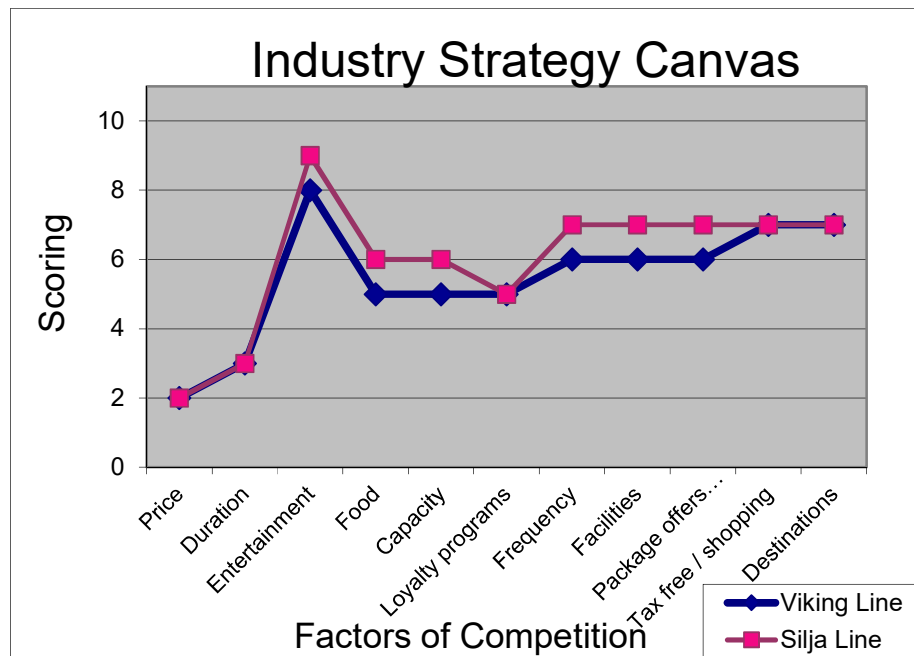
By looking in the strategy canvas, it is visible that companies need to find the way to produces unique value through differentiation; by choosing different activities; by creating more value to customers.

Twitter's user activity looks very low and similar for both companies for the first monitoring months. First impression is that no needs for data aggregation and analysis. Below char shows tweets and retweets about Silja Line.

Tweets vs Retweet
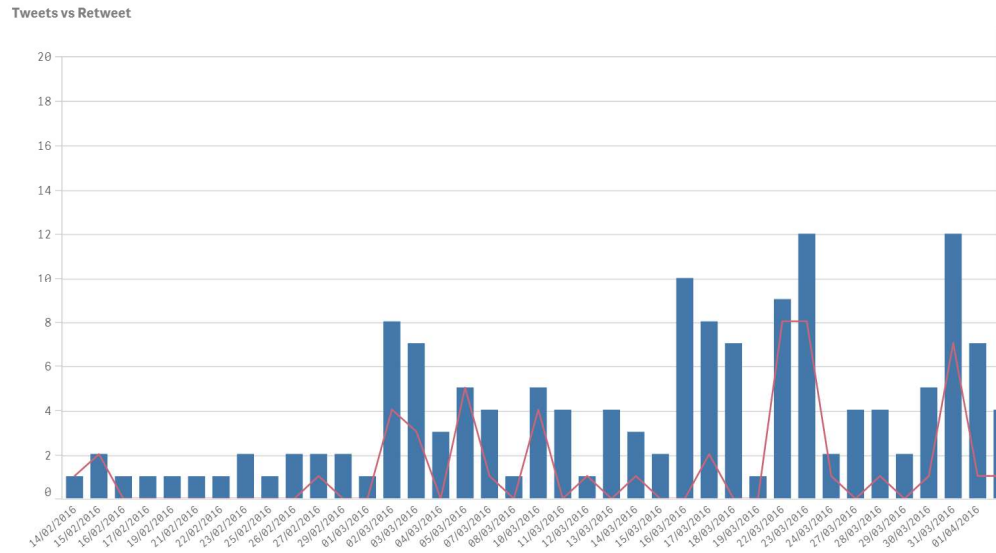
The next char shows tweets and retweets about Viking Line.

Tweets vs Retweet
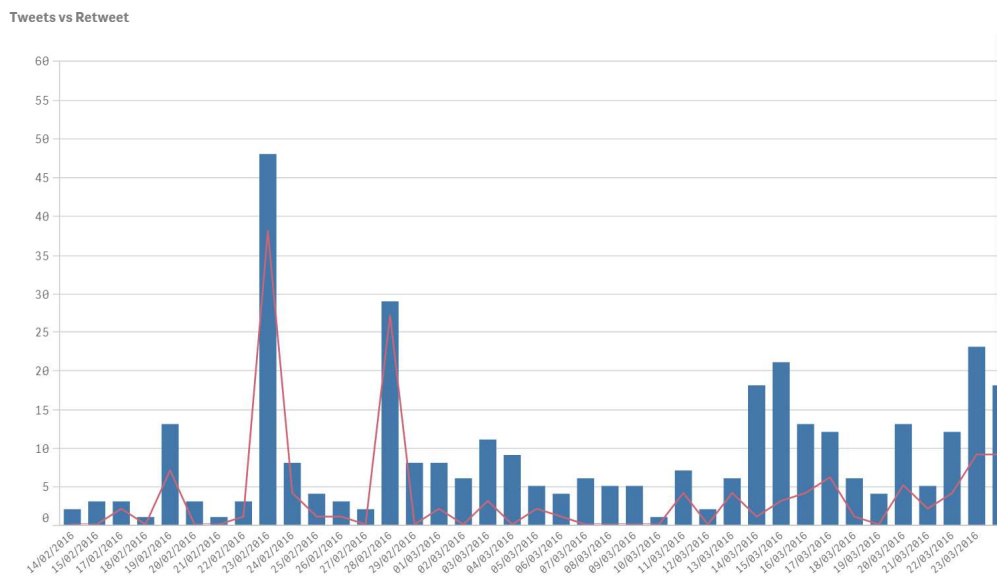
But at the end of the spring, there was big news regarding norovirus epidemic into Silja Symphony cruise ship. The epidemic took a few weeks. It started on 29.5.2016. Yle (2016) wrote that "over the past week, an estimated 150 people on the cruise ship Silja Symphony have fallen ill from what health authorities confirm as norovirus". Regarding the epidemic continuous news coverage during the best tourist peak season, there was not any impact on twitter's user activity or any opinion feedback until

06.06.2016.

But user activity, started from 6.6.2016 was not related to the epidemic. It was regarding Robin's gig in the ship. There are 100 tweets was found that include "Robin" word and only six that include "noro" word from the whole collected data of the company.

The charts below show the tweets, which included word "Robin".

The chars below show the tweets, which included word "Noro".

It seems that a celebrity definitely brings visibility to a company. Additionally, from this example, it is visible that Silja Line has strong brand with good reputation. Because, constant negative news published by traditional media regarding the company did not bring significant negative conversation into twitter.

The confirmation of the discoveries may be found by looking to the stock chart. The epidemic news did not make any negative impact into the stock price. Furthermore, it seems that the Robiner gig event influence positively into stock price.

Figure 16: Stock price for symbol TAL1T.TL from 13.5.2016 to 15.6.2016

By looking in Viking Line twitter collected data from the monitoring time period, it is observed that there are three days with twitter user activity

picks. They are 5.5.2016, 12.5.2016 and 9.6.2016
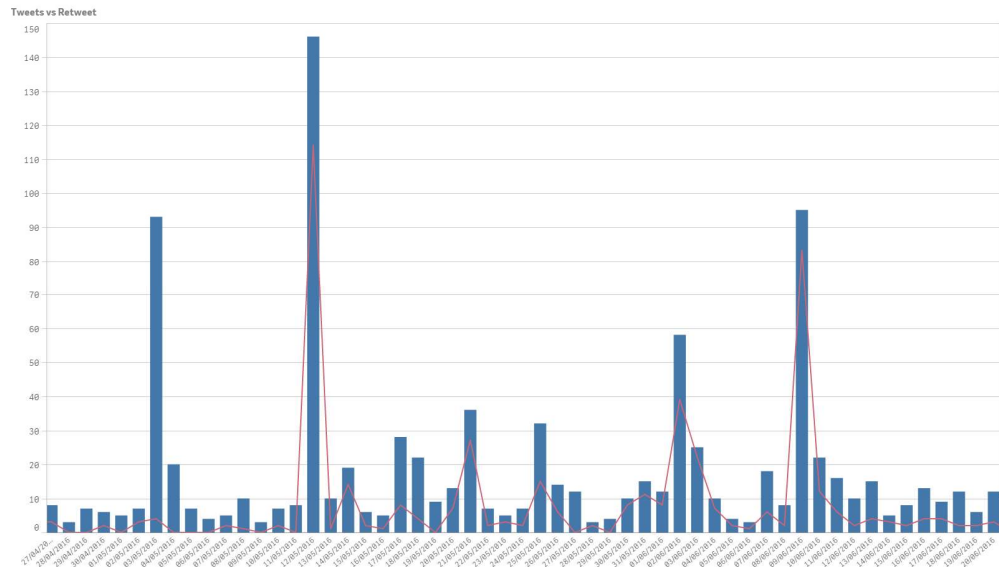
On 5.5.2016, there are 93 tweets and only 4 retweets. They are mostly related to the news regarding to the new Viking line plans to launching 141-day world cruise.
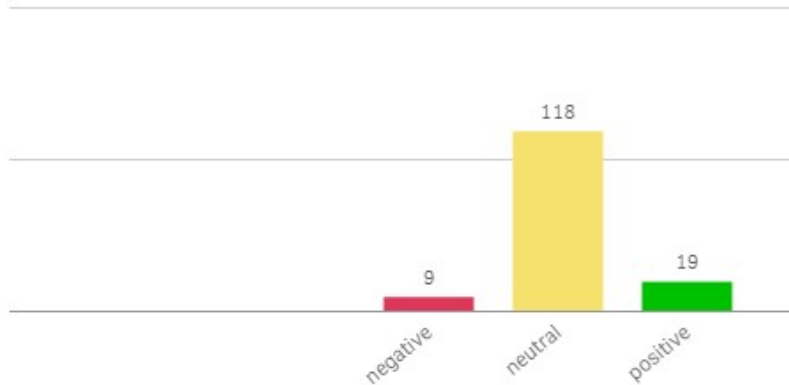
Tweets

nickname	Score	tweet
_CruiseNews	0,00	New Viking line plans epic, 141-day world cruise https://t.co/tgBXmwFT3g
_travelicious	0,00	New Viking line plans epic, 141-day world cruise: The sailing will be the first of its kind for Viking Ocean ... https://t.co/SBtFWt5OMW
4UET	0,00	New Viking line plans epic, 141-day world cruise: The sailing will be the first of its kind for Viking Ocean ... https://t.co/ErEKPHmS57
140CruiseNews	0,00	New Viking line plans epic, 141-day world cruise - USA TODAY https://t.co/0wUSnfmVW3 #cruise #cruisenews
adisonwarren	0,00	RT @besthotelstay: New Viking line plans epic, 141-day world cruise https://t.co/vmkL5zLGSC https://t.co/LMbPJIlJOU
ADMIRANDA3	0,00	RT @besthotelstay: New Viking line plans epic, 141-day world cruise https://t.co/vmkL5zLGSC https://t.co/LMbPJIlJOU
alexxamartinez	0,00	New Viking line plans epic, 141-day world cruise: The sailing will be the first of its kind for Viking Ocean ... https://t.co/LaKkxbnPsN
annabelle_dirk	0,00	New Viking line plans epic, 141-day world cruise https://t.co/NOekawJgSY https://t.co/PTKX1MAEYe
Baileyavis	0,00	New Viking line plans epic, 141-day world cruise: The sailing will be the first of its kind for Viking Ocean Cruises

On 12.5.2016, there are 146 tweets and 114 retweets. Most of the twitter's' user activity related to put into service the digital travel experience application. (Viking Line, 2015) The peak in twitter activity, with significant amounts of retweets reflected into stock price curve positively, regards a few messy feeling from users.

Adopted from Twitter.

Sentiment by count

Sentiments representation in 12.5.2016.

Figure 18: Stock curve for symbol VIK1V.HE from 5.5.2016 to 15.6.2016

But after 9.6.2016, we can see drop in the stock price that happened at the same time with a lot of twitter user activity (95 tweets and 83 retweets). 82,1 % of the tweets are negative. The stock price fall continues until 17.6.2016.

The print screen below shows the entire overview of Viking Lines' data on 9.6.2016. It includes a lot of negative comments related to the broken workers' rights issue. Strong emotions regarding company activity signify ratio of retweets and tweets (83 vs. 95) that presented in Tweets vs Retweet chart.
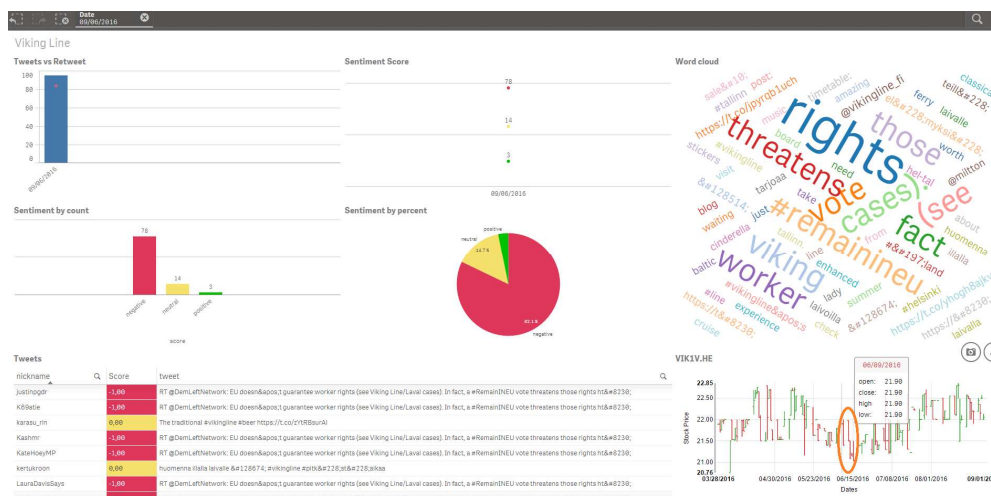
During the loss period, Viking Line decided to re-flag a ship to operate under a "flag of convenience" in Estonia or Norway. It was done in order to avoid collective agreements with Finnish trade unions and cut jobs and terms and conditions. (Thompsons solicitors, 2008)

Comparison of overall negative and positive rations of the companies may be used to find out the users' likes between the companies. The same analysis can be extended over multiple time periods for trend analysis.

Silja Line

- Negative ratio or percentage of negative conversations over total conversation.

Negative ratio = 48 / 848 = 0,06 %

- Positive ratio or percentage of positive conversations over total conversations
  Positive ratio = 104 / 848 = 0,12 %

Viking Line

- Negative ratio or percentage of negative conversations over total conversation
  Negative ratio = 179 / 1659 = 0,11 %
- Positive ratio or percentage of positive conversations over total conversations
  Positive ratio = 150 / 1659 = 0,09 %

Silja Line created more positive impact on customer's experience for the data extract period then Viking Line. Because positive ratio is higher and negative ratio is lower for Silja Line compared to Viking Line. The findings confirm trends of the stock price development over the reference date collection period.

3.2.2 Crisis control

Over the past few years, fraud companies' scandals have become a public, such as Volkswagen emissions scandal in September 2015. (Valentine, 2015) Next cheating revelation did not have to wait for a long time. On 26 February, Nokian Tyres has admitted that they created custom, high-quality versions of tyres that were used during tests. Those were not otherwise available to the public in the exact same form. They did so in order to get better test results and hence positive publicity for its products for years. (Yle, 2016) The scandal leds to a 9.6 percent fall in Nokian's shares at one point on Friday. But already in the next day, share price start to recover. The same trend line is visible in the stakeholders' twitter conversations. It seems that strong brand helps company to recover the reputation from this scandal.

Figure 19: Stock curve for symbol NRE1V.HE from 14.2.2016 to 15.3.2016
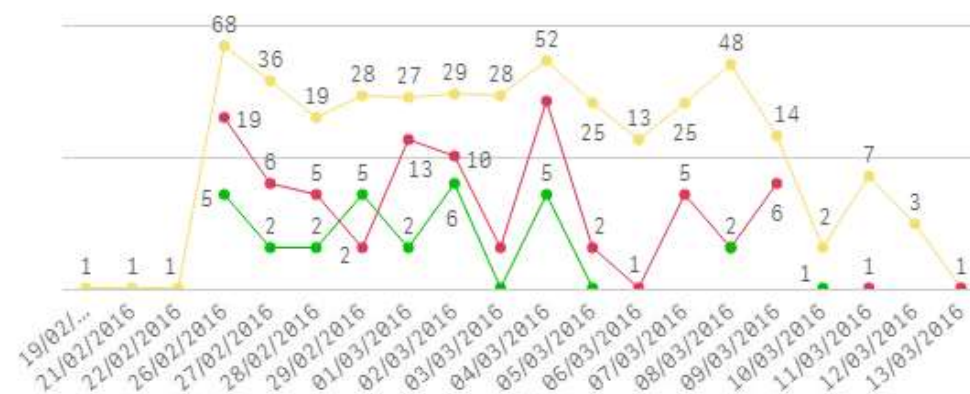
On March first, there is increase in the negative conversation activity. At the same time, stock value starts to declines again until 10 March. The reason for this is the report that the CEO sold all of its ownership in Nokian Tyres before the disclosure of the manipulation ring. (Pietiläinen, 2016) On March fourth, Helsinki Sanomat newspaper told that there are nine from 12 members of the Executive Board who gave up options valued at 600 000 euros last autumn. (Pietiläinen, 2016)

This scandal left bad damage to the company's reputation and hurt notably brand loyalty. Because it tooks approximately two weeks for getting away from negative conversation in social media. Additionaly, the tweets about

company generated a lot of the retweets. It shows that the tweet text brings emotions and desire participates in the discussion.

Tweets vs Retweet
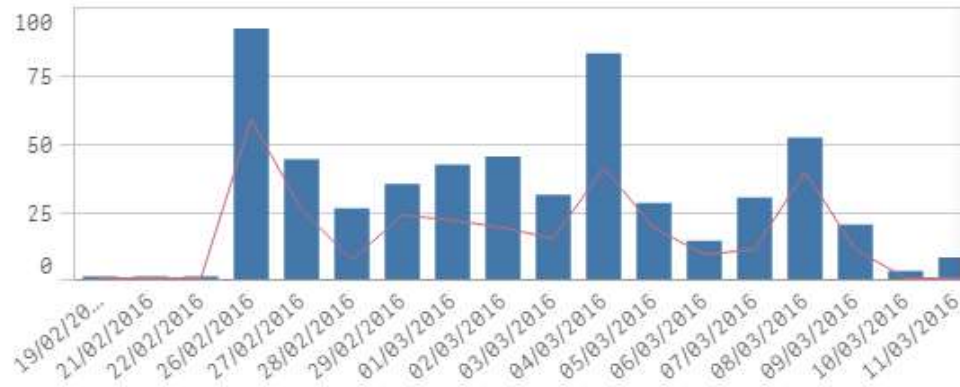
This conclusion confirms the trend of stock price and the NokianTyres company's interim report 2016, Q1. The report told that net sales dropped -1,9% compared to the same period in 2015. (NokianTyres 2016)

As a reference, the author used the Wärtsilä manipulation on ship engine fuel tests case on March 7. Wartsila admitted that workers, at one of its businesses in Italy have manipulated a small number of fuel consumption tests for its ship engines. They said an internal audit revealed that 2 percent of its ship engine deliveries may have been affected by the manipulations. Jornal writes "While the wrongdoing appeared modest in scale, this kind of violation has been in sharp focus since the Volkswagen AG emissions scandal broke last year." The fact is that fuel costs are a key determinant of profitability for shippers. Shares in Wartsila fell 2.9 %. (WSJ 2016)
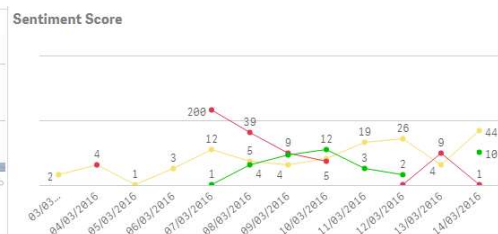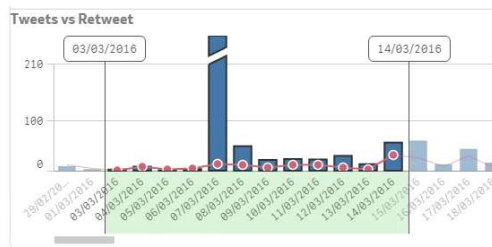
The grievous threat may be analysed by looking in the stakeholders' social media activity. There are a lot of the tweets but only a few retweets on March 7. At the next day, the conversation drops from 213 to 48 tweets. Additionally, the sentiment score shows that negative tweets expiried three days later, on March 10. So, by this analysis it is visible that this news did not cause the big damage to the company reputation or brand.

The finding confirms that the stock price drop recovered from the harm very quickly, in two days. The company's interim reports, from July 2016 told about stable orders and sales in challenging markets. It says orders increased by 3 %. (Wärtsilä 2016)

Figure 20: Stock price for WRT1V.HE from 3.3.2016 to 21.3.2016

3.2.3   Panama papers

Next, there was a case of embezzlement this year, the so called panama papers, which was exposed on 4.4.2016. Nordea is one of the biggest banks in Scandinavia. It has been accused of offering clients complex structures to hide their assets to dodge taxes. (Reuters 2016) First weeks after scandal, Nordea pretended that nothing happened. The company representative said that there was no evidence Nordea had done "anything illegal, though the deluge of criticism from media and

shareholders had forced the bank to respond to allegations of unethical conduct." (Schwartzkopff, 2016) But this strategy definitely did not help to take the crisis into control. There were a lot of stakeholders' twitter activities related to Nordea Company with dominant of negative conversation over positive.
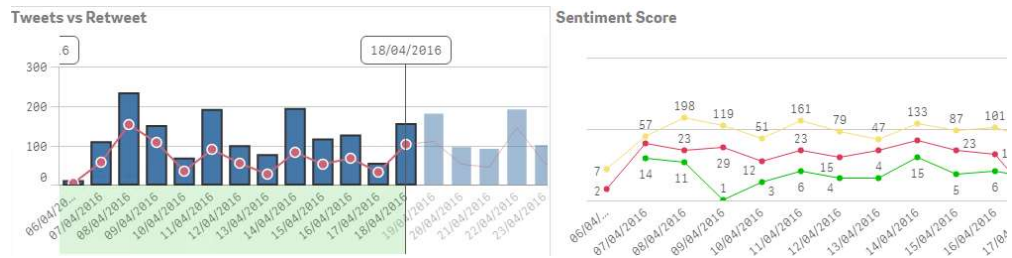
Huokaus. Kriisiviestintä laahaa. #viestintä #panamapapers #nordea

DNB JA Nordea ovat tienneet viikkokausia, että niiden yhtey-det Panamaan paljastuvat. Aikaa olisi ollut kunnon toi-menpideohjelman ja viestintäsuunnitelman rakentamiseen. Tahtoa ...sti puuttui.

...pson jäi kiinni törttöilys-...siaa vaimolleen: "Mutta ...na, ettet koskaan saa tie-

5.55 - 8. huhtikuuta 2016

Adopted from Twitter.

On April 27, Chief Executive Officer Casper von Koskull finally admitted the bank's alleged role in helping rich clients hide their wealth. To save the company reputation after the revealed, Nordea Bank decided "engaging with" customers as much as possible. Nordea admitted that profit lunged 28 percent last quarter at the same time with revenue slip. (Magnusson & Schwartzkopff, 2016) At the end of the April, Nordea started to do corrective actions. Their results may track by sentiment analysis char.

The below chart shows the positive tweets volume over the time that helps to track the corrective action from Nordea site.

Sentiment Score

On April 25, Nordea announces plan to launch a new Crowdfunding equity platform. This news has got a positive feedback from stakeholders.

Adopted from Twitter.

On May 12, it was press release regarding opening a Nordea Accelerator programme applications that will run for 12 weeks starting September 2016. The programm helps startups get a new chance to accelerate their business.

nickname	Score	tweet
emeliearvid	1,00	Good news for #fintech #startups from Nordea! New #NordeaAccelerator programme now open for applications: https://t.co/M2cYDMBZyF
frkVivi	2,00	@inayulo great👍
Please do read info and apply here https://t.co/yP9RcZv80p. No emails :) @LivFreihow @Nordea #NordeaAccelerator

The most notable positive effect was made Nordea Masters 2016 event that represent a lot of positive twitter activity during June 6 and 7.

Tweets

| nickname	Score	tweet
Golfalot	1,00	Nordea Masters winner Matt Fitzpatrick uses two @PINGTourEurope G fairway woods https://t.co/QjKZ4DcqvA https://t.co/x5A3ZETkmI
golfcentraldoc	1,00	Stats Category Winners: Who Did What Best At The 2016 Memorial And Nordea Masters: Always nice on a Monday mo... https://t.co/99PNtYD1dE
GolfCourier	1,00	#Golf #GolfNews Matt Fitzpatrick leads Nordea Masters in Sweden by five shots https://t.co/BavLGT2Nqn
GolfDigestSA	1,00	WEEKEND WINNERS
M Fitzpatrick (Nordea Masters), W McGirt (Memorial), Jaco Ahlers (Zambia), A Nordqvist (ShopRite). https://t.co/bAYTFh8AaX

Despite the correction actions from Nordea side, it is clear that negative twitter activity overflew at the monitoring period, especially at the middle of the summer.

Sentiment Score

On June 21 there are the visible pick of negative activity. It is related to speculation that Nordea does have enough capital to fulfil all capital

requirements.
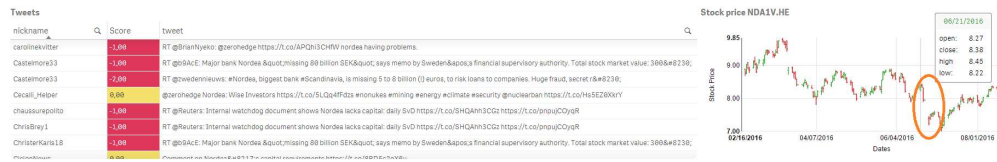
As a result, Nordea shares were down from 8,39 on 21 June to 7,37 euros on June 28.

Figure 21: Stock price for NDA1V.HE from 3.3.2016 to 6.7.2016

The Metso panama papers case presented below as well. It has been done to compare previous, Nordean and Metso panama papers cases. The comparison demonstrated the difference in sentiment scores and the reaction of the stakeholders to the embezzlement.

On April 11, Metso Corporation has press release about getting information from media that the company are connected to "panama papers". Metso immediately initiate an internal audit process regarding this case with the independent external auditing company PwC. The audit finds that only one employee has acted against the company instructions, outside the scope of company operations. The employment contract of this employee was subsequently terminated. (Metso, 2016) At the same day, Metso present the innovation programs at Metso Aggregates Academy during Bauma. That was widely reported by media and the news got a lot

of positive feedback from stakeholders on April 12. (Nasdaq, 2016) Metso avoided crisis by immediate acting on the negative news and brining positive news to a public. This can be easily tracked by sentiment analysis.

The print screen below shows the entire overview of Metso's data for period from 9.4.2016 to 19.4.2016. Negative tweets mostly stop in a two days. Additionally, positive tweets dominate over negative tweets in a next day after "panama papers" has been revealed.

"Panama papers" scandal has no negative impact into stock price. It was positive stock rises day.

Figure 22: Stock price for MEO1V.HE from 6.4.2016 to 22.4.2016

4   CONCLUSIONS

This study attempts to identify the specific opportunities and challenges that offer the social media analysis to companies. The aim of the research was to find the way to additional business value creation. The author presented an effective and practical approach to monitor reputation polarity by using twitter activity. That helps the companies to see customers' response on a product or service, to react quickly to the challenges situations.

Analysis of social media helps companies to get to know their customers and how they view the company. Word cloud helps to understand what the most common impressions of company's brand really are. By monitoring social media activity, companies have a higher possibility to improve business performance.

4.1   Conclusion of the research objectives

The main question of the research is how may social media be utilized in strong reputation construction?

An online activity shows stakeholders' response on company operations. By following their activity in social media, the companies may easily identify issues that may impact the company reputation. The stakeholders' response can be monitored by collecting and analysing the social media comments. Monitoring allows the company to quickly revise status of online activities and make quick adjustment while they are still manageable. It is allowing to react quickly on changes and to monitor, is there are correction movement have effect or not.

The second objective was to find out how may online opinion data mining be utilized in reputation building?

It is very important to identify the level of customers' satisfaction regarding a company operations. Recognition of an emotional value behind social media posts enables any business to draw deeper insight through the

analysis of unstructured data. The company may gain a trust and loyalty of the customers by following their online activities in a timely manner. The real time "survey" helps to determine an attitude, opinion and emotion expressed by the customers iana online comments. By sentiment analysis the company may know the emotions that stakeholders feel to get better understand their behavior.

How do online customers' opinion impact reputation and business performance?

The research shows that social media monitoring helps to improve a chance for any company business success. Because there is a direct connection between business performance, like stock price or sales and stakeholders' opinion expressed in social media. Verifiably, people have a high-speed and power voice channel thru social media. Social media comments play an important role in the consumer decision making process.

Why it is needed to monitor and measure opinions in the first place?

Since social media is a public, company operations are completely transparent. A customers' feedback regarding a product/service or uncover needs are visible for other users and may influence their purchase decisions. The companies forced to dig deeper into those conversations, particularly into the negative one. It may be a weak signal to make a correction movement. Even one negative comment sometimes can quickly go like viral and impair the company reputation.

Is there are correlation between online opinions and reputation?

In the Nordea case, the author introduced the strong relation ships between online opinions and reputation. Large-scale word-of-mouth networks need to be taken seriously. A technology need to be engaged for assisting building trust and fostering cooperation with stakeholders in online.

How does sentiment analysis impact the decision making outcomes?

A company needs to benchmark, measure and monitor activities to see how well they are performed. Strategic evaluation and control is a significant assessment process that determine whether the chosen activities is achieving the organization's goals.

How valuable are these insights for business intelligence?

Stakeholders' opinion is important because it provides business owners with metrics that they can use to manage and improve their businesses. Customer's word of mouth is the best indicator of consumer lifetime, repurchases intentions and loyalty.

How can a company use social media data for performance measurement?

Key performance indicators (KPI's), presented by the author can be used as quantitative indicators of critical success factors for an organization. They help to understand how company is doing against objectives. The KPIs are an actionable scorecard that intends to keep company's strategy on track.

4.2   Recomendations

The outcome of this study shows that customers activity in social media channel is a useful source of information for any business, despite an industry they operation in.

According to Howson and Schelte (2016) webinar, an operational data warehouse has no longer sufficient to get data inside. The traditional data warehouse needs to be combined with a public and social media data to get leading indicator. Nowadays, companies do not use social data in the holistic way. The companies maybe have some social analysis on the

side. But for a complete analytics approach it is not sufficient anymore.

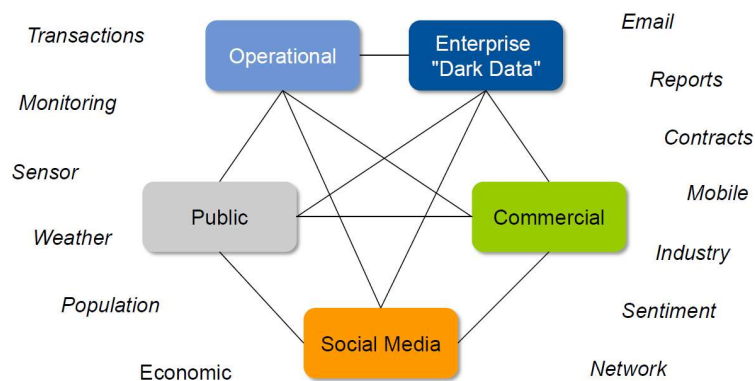## Analytic Businesses are Concerned with More than Just Transaction Data

Figure 23: Analytic businesses are concerned with more than just transaction data (Howson & Schelte, 2016)

The author would like to recommend combining a sales performance data with a public, social media data to get better insight into a company performance.

The question is how to utilize a data into value creation for a business or how to extract knowledge from the data via technologies that incorporate the below principles. The next steps designed to assist in identifying business problems from the data perspective:

1. The first step is to evaluate business strategy. It can be divided into: Growth, Cost reduction or Risk reduction strategies.
2. The next step is to agree strategic objectives and priorities by understand the real business questions. It will help to get the answer they really want to know.
3. Create a data analysis plan and defined the source systems for data.
4. Collect and validate the data based on step 3.
5. Analyse the data and extract the data insights.
6. Evaluate the insights and present for the business decision.

4.2.1 Lessons learned

The web 2.0 era has made reputatioal and brand managenig prosesses more elaborate, complex and meaningful that ever before. The online reputation becomes more and more the result of subjective opinions shared by users in social media networks. (Aula, 2010). Moreover, information and opinions can spread through online communities rapidly without any control from a company. According to Gaines-Ross (2010), most of the companies are slow moving but the ability of quick response to reputational changes is a necessary condition for the achievement of good reputation.

Emergence of online and open communication environments has created a need for a new approach to monitor, measures and manages online customers' activities. Ignoring online voices of customers isn't an option any more. In the past, companies used not responding to the customers' concerns. But now, the companies need to shows to the customers that those are taken seriously. The companies have to learn not only communicate effectively in the social media. But more important is listen their customers and respond in the way that aligns with both company stratagy and customer expectations.

Social media as a voice of customers gives a lot of opportunities for a company to achieve a better profitability. It helps to monitor and improve the company reputation and brand equity in near real-time rate. Social media has offered a host for new platforms for the companies to build relationships, foster engagement and better understand the needs of the customers. However, while social media has broken traditional boundaries by creating a sense of closeness and transparency between the companies and the customers, it has also made the companies more vulnerable to potentially damaging situations.

4.3   Validity and reliability of research

Key indicators of a quality of research are reliability and validity of measures. In general, reliability of the research represents an ability to obtain the same or at least very similar result in repeated measurements on the same selection. The assumption is that the conclusions have remained the same between researche. On the other hand, validity represents a confidence that the research really measured the concept that was intended to measure. (Vehkalahti, 2008)

Maylor and Blackmon (2005) continue listing the requirements for good research. Additionally, to the validity and reliability, research need to be generalisable and creditable. They mean by generalisability that findings may be applied widely, beyond the research data set. Credibility represents veracity of the research results.

According to QMSS, one way to assess reliability of the research is to exhibit to meet three criteria of reliability.

1. Test-retest reliability ensures that the result will be the same if the researcher uses multiple times the same measurement tool, asks the same question and follows the same research procedures. It is show that the results are consistent.
2. Inter-item reliability can be used in cases where multiple items are used to measure a single concept. In this case a research needs to be design so the questions and the measures should be associated with each other.
3. Interobserver reliability ensures that different research using the same measure will get the equivalent results. (QMSS, 2016)

QMSS writs, that a valid measure should satisfy next criterias.

1. Face validity evaluates when a research measures the concept it is intended to measure.
2. Content assess is the validity of the concerns in which a measure adequately represents all of the concepts.

3. Construct validity test is the dimension that the research measures the essential criterias available. (QMSS, 2016)

The research can not be reliable without being valid. Because of this, the author wants to discuss validity aspect of the research first.

Clear questions conducted objectives and goals that were answered during the research. The questions' answering was supported by associated KPI, related literature review and of existing social media research in the same subjects' area.

To ensure the research reliability, the author paid close attention to the data gathering and analysis quality results. For this, data cleaning and mining was performed to ensure that the method of data gathering leads to consistent results. (Explain future in paragraph 3.1 Research approach to data source.) During the research, the measures of the results were compared with existing research results and data available. It was done for purpose to find out that results are comparable for many aspects. Because the results are similar then it is likely that the method of data gathering is reliable. Reliability represents the internal consistency of measurement instruments because measurement was performed by the same application over all cases data. That has deen tested during application developments by using test cases, e.g. validity of sentiment analysis evaluated by comparing sentiment scores for specific comments to personal assessment of the text opinion.

Data sampling was performed widely and for a long period of time that ensure the data consistency and reliability as well. Detaild information may be found in paragraph 1.2 Research objectives and scope.

However, the several issues may affect the accuracy of data collected that has explained in paragraph 2.3.3 How sentiment analysis works.

The research main goal is to prompt any business to listen to customer's voice by using new channel in general. Social media analysis investigated

as an easy source of the real-time customer's data available for everyone. So, the research meets generalisability criteria as well.

In conclusion, the research seems to be reliable and valid in terms of content and usability. However, enhancing data source into other social media application (like Facbook) could improve the result.

4.4   Proposals for further actions and research

Despite the importance of this topic for both academics and practitioners, it is a very few studies investigate the strategic and operational mechanism of utilizing the social media. It is very valuable to research the all possible ways of creation real benefits of a company by using public data.

One opportunity for a business, in social media utilization may be co-creation with customers in research and development area. It would be interesting to investigate how the word cloud could be utilized in customer participation in a new product development, in co-creation with a customer; get the ideas from the customers. It would be very useful to find out additional opportunities from social media utilisation in other business development aspects as well.

The type of correlation between online opinion and stock market should also be an area of focus. This area needs to be more investigate because there are many contradicted studies available. So, it worth to find the truth what sort of correlation between the stock price and online opinion. But more important is to find how this correlation needs to be utilized in a best possible way.

Risk management and reduction is the one of the important factor in a succsefull business stratagy. Regardless of the industry, every company faces its own set of risks in some point of time. So it would be very beneficial to find out social media integration benefits into crisis management and into crisis communication strategy. How to create a successful crisis online communication strategy. How catch this moment when the crisis is developing on social media. This is the key to effectively

react on it. Having a real-time overview of the situation allows companies to keep ahead of a game. Because being able to anticipate potential threats are crucial, to ensure that a company is well positioned to manage crisis situations.

One of the very interesting areas, that attract a lot of an attantion is social media utilisation in trends' forecasting and political events, like prediction of the election results or United Kingdom voted for the resignation of the European Union.

The author fined that celebrity brings a lot of attention from the social media users. So, utilisation of this opportunity in brand awareness needs to be a subject for a research as well.

REFERENCES

Written references

Ackoff, R. L., (1999). *Re-creating the corporation: A design of organizations for the 21st century.* OUP Oxford. 159-163.

Aula, P. 2010, 'Social media, reputation risk and ambient publicity management', *Strategy and Leadership*, vol.38(6), pp. 43-49.

Aula, P. & Heinonen, J., (2016) *The reputable firm*. Springer. 4, 20, 23-24

Barnett, L., Jermier, M. & Lafferty, B., 2006, 'Corporate Reputation: The Definitional Landscape', *Corporate Reputation Review*, Vol. 9, No. 1, pp. 26–38

Beesley, L. 2005, 'The management of emotion in collaborative tourism research settings', *Tourism Management*, vol. 26(2), no. 2, pp. 261-275.

Blanchard, O., (2011), *Social media ROI: managing and measuring social media efforts in your organization*, Que, Indianapolis (IN). 18-24, 26.

Bunting, M. & Lipski, R. (2001), 'Drowned out? Rethinking corporate reputation management for the internet', Journal of Communication Management, Vol. 5 No. 2, pp. 170-178.

Chan, K. & Mauborgne, R. 2002, Charting Your Company's Future, Harvard Business Review, June 2002, Vol. 80 (6), 76-83, 153

Fernández-Gavilanes, M., et al., 2016,'Unsupervised method for sentiment analysis in online texts', *Expert Systems with Applications*, vol.58. 58

Fombrun, C., (1996) Reputation: Realizing Value from the Corporate Image. Hardvard Business School Press. 68

Gaines-Ross, L. 2010, Reputation Warfare, Harvard Business Review December 2010. 70-76.

Gould, P. (1981) *Letting the data speak for themselves*. Annals of the Association of American Geographers. 166.

Graeme, M., Ronald, J., & Cary, L., (2012) *Corporate Reputation*, Gower. 3

Hammond, J., (2008), *Branding your business: promoting your business, attracting customers and standing out in the market place*. Kogan Page, Longon. 14, 21, 31.

Han, J., Pei, J., & Kamber, M. (2006), *The Morgan Kaufmann Series in Data Management Systems: Data Mining, Southeast Asia Edition: Concepts and Techniques (2),* Morgan Kaufmann. 6-7.

Hennig-Thurau, T., Hofacker, C. F., & Bloching, B., 2013, 'Marketing the pinball way: understanding how social media change the generation of value for consumers and companies', *Journal of Interactive Marketing*, vol.27, no.4, pp. 237-241.

Hu, M., & Liu, B. (2004) *Mining and summarizing customer reviews. In Proc. of the 10th acm sigkdd international conference on knowledge discovery and data mining*, New York, USA: ACM. pp. 168–177.

Johnson, G., Scholes, K., & Whittington, R., (2008), *Exploring Corporate Strategy*. London: Prentice Hall. 9.

Kaur, A., & Gupta, V., 2013, 'A Survey on Sentiment Analysis and Opinion Mining Techniques', *Journal of Emerging Technologies in Web Intelligence*, vol 5, no 4, pp. 367-371.

Kim C., Mauborgne R. 2005, Blue Ocean Strategy, Harvard Business Review May 2014, 77-84.

Kohli, C., Suri, R., & Kapoor, A., 2015, 'Will social media kill branding?' *Business Horizons*, vol. 58(1), pp. 35-44

Kramer. L., (2010), How French Innovators Are Putting the Social Back in Social Netwroking, Harvard Business Review October 2010. 121-124.

Laaksonen, S-M., et al. (2012), *Digital Reputation; Characterizing and parameterizing reputation, reputation risks and the impact of digital publicity*, Helsinki: Media and Communication Studies Research Reports.

Lovett, J., (2011) *Secrets : Social Media Metrics Secrets*, Wiley. 4-6, 349-350.

Lui, B. (2012) *Sentiment Analysis and Opinion Mining*, Morgan&Claypool. 2, 13-14.

Maylor, H. & Blackmon, K. 2005, *Researching business and management,* Palgrave Macmillan, Basingstoke. 362-381

McElheran, K., Brynjolfsson, B., 2016, The Rise of Data-Driven Decision Making Is Real but Uneven, Harvard Business Review February 2003. 2-5

McGrath R. 2010, 'Business Models: A Discovery Driven Approach', *Long Range Planning*, Volume 43(2), pp. 247-261.

McGrath R., MacMillan I. 1995, Discovery-Driven Planning, Harvard Business Review July-August 1995. 44-54.

Minelli, M., Chambers, M. & Dhiraj, A. 2013, *Big Data, Big Analytics*, John Wiley & Sons, Inc, New York. 37-38, 40.

Mohanty, S., Jagadeesh. M., & Srivatsa, H., (2013) *Big Data Imperatives*. Apress. 1, 9-10.

Oracle 2012, Why Customer 'Satisfaction' is No Longer Good Enough.

Porter, M. 1996. Strategy for Business' What Is Strategy?, Harvard Business Review November-December 1996. 4-13.

Porter, M. 2008, The five competitive forces that shape strategy, Harvard Business Review January 2008. 79-82.

Poynter, R. 2010. The Handbook of Online and Social Media Research: Tools and Techniques for Market Researchers (1), John Wiley & Sons.

Russell, G., Sean, C., (2014) *Big Data-Driven Business: How to Use Big Data to Win Customers, Beat Competitors, and Boost Profits*. Wiley. 17-21, 24, 78, 117.

Schnietz, K. E., and Epstein, M. J., 2005, 'Exploring the financial value of a reputation for corporate social responsibility during a crisis', *Corporate Reputation Review*, vol.7(4), pp. 327

Smythe, J., Dorward, C. & Reback, J. (1992) *Corporate Reputation: Managing the New Strategic Asset*, London: Century Business Ltd. 7

Vehkalahti, K. 2008. Kyselytutkimuksen mittarit ja menetelmät. Vammalan Kirjapaino Oy, Vammala. 41.

Wartick, S. 1991, 'The relationship between intense media exposure and change in corporate reputation, *Business and Society*. 33-49

Weigelt, K. & Camerer, C. 1988, 'Reputation and corporate strategy: A review of recent theory and applications', Strategic Management Journal, 9.

Yoon, E., Guffey, H. & Kijewski, V. 1993, 'The effects of information and company reputation on intention to buy a business service', *Journal of Business Research*. 27

Zook C., Allen J. 2009, Growth outside the core, Harvard Business Review December 2003. 67-73.

Zook, C. 2007, Finding Your Next Core Business, Harvard Business Review April 2007, 68-69, 71, 74.

Laws, statutes, decrees, committee reports and standars

European Union. (2010) *European textbook on ethics in research*. Luxembuourg: Luxembuourg Publication Office of the European Commission, viewed 12 January 2016, https://ec.europa.eu/research/science-society/document_library/pdf_06/textbook-on-ethics-report_en.pdf

Manyika, J., et al. (2011) *Big data: The next frontier for innovation, competition, and productivity*, McKinsey. 4-8

Electronic Sources

Brynjolfsson E., Hitt L., & Kim H. 2011, 'Strength in numbers: How does data-driven decision making affect firm performance?', SSRN working paper, viewed 10 Februry 2016, http://ssrn.com/abstract=1819486

Chaffey, D. 2015, 'Global social media research summary 2015', Smart Insight, viewed12 December 2015, http://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/

Data and data management glossary 2016, DDDM, viewed 10 Februry 2016, http://whatis.techtarget.com/definition/data-driven-decision-management-DDDM

Duxbury, C. 2016, 'Engine Maker Wartsila Admits Some Fuel Consumption Tests Manipulated', The wall street journal, viewed 3 July 2016, http://www.wsj.com/articles/engine-maker-wartsila-admits-some-fuel-consumption-tests-manipulated-1457359468

Dwyer, J. 2010, 'A Clothing Clearance Where More Than Just the Prices Have Been Slashed', New York Times, viewed 1 July 2016, http://www.nytimes.com/2010/01/06/nyregion/06about.html?_r=0

European Commission 2012, Europeans and their languages, viewed 12 May 2016,
http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf

European Commission Directorate-General for Research, European textbook on ethics in research, viewed 12 May 2016,
https://ec.europa.eu/research/science-society/document_library/pdf_06/textbook-on-ethics-report_en.pdf

Gualtieri, M. 2012, 'The Pragmatic Definition of Big Data', Forrester, viewed 12 May 2016, http://blogs.forrester.com/mike_gualtieri/12-12-05-the_pragmatic_definition_of_big_data

Howson, C. & Schelte, R. 2016, 'BI and Analytics Trends for Digital Business', Garner, viewed 12 January 2016, http://www.gartner.com//it/content/3237400/3237417/april_14_bi_and_analytics_trends_howson_schulte.pdf?userId=84695940

Keane, M. 2010, 'H&M's Trashgate: A lesson in social media', Econsultancy, viewed 1 July 2016, https://econsultancy.com/blog/5221-h-m-s-trashgate-a-lesson-in-social-media

Koho, S. 2015, 'Tutkija Veetin Facebook-mainonnasta: "Kääntyy pahasti itseään vastaan"', Markkinointi&Mainonta, viewed 29 January 2016, http://www.marmai.fi/uutiset/tutkija-veetin-facebook-mainonnasta-kaantyy-pahasti-itseaan-vastaan-6295777

Laney, L. 2012, 'Deja VVVu: Others Claiming Gartner's Construct for Big Data', Gartner, viewed 12 May 2016, http://blogs.gartner.com/doug-laney/deja-vvvue-others-claiming-gartners-volume-velocity-variety-construct-for-big-data/

Magnusson, N. & Schwartzkopff, F. 2016, 'Leaks ',Bloomberg, viewed 10 July 2016, http://www.bloomberg.com/news/articles/2016-04-27/nordea-profit-sank-28-last-quarter-as-bank-s-revenue-declined

Metso 2016, Annual Reports 2015, viewed, viewed 10 July 2016, http://annualreportmetso.com/wp-content/uploads/2016/03/Metso_Annual_Review_2015.pdf

Metso 2016, Metso comments on information presented in the media, viewed 10 July 2016, http://www.metso.com/news/2016/4/metso-comments-on-information-presented-in-the-media/

Mishne G. & Glance N. 2006, 'Predicting movie sales from blogger sentiment', AAAI, viewed 21 July 2016,

https://www.aaai.org/Papers/Symposia/Spring/2006/SS-06-03/SS06-03-030.pdf

Nasdaq 2016, Metso shapes the industry by introducing easier ways of buying, maintaining and operating aggregates equipment, viewed 10 July 2016, http://www.nasdaq.com/press-release/metso-shapes-the-industry-by-introducing-easier-ways-of-buying-maintaining-and-operating-aggregate-20160412-00091

Nemschoff, M., 2013, 'Social Media Marketing: How Big Data is Changing Everything', CMS Wire, viewed 12 May 2016, http://www.cmswire.com/cms/customer-experience/social-media-marketing-how-big-data-is-changing-everything-022488.php

Nexgate 2013, State of Social Media Spam, viewed 11 July 2016, http://nexgate.com/wp-content/uploads/2013/09/Nexgate-2013-State-of-Social-Media-Spam-Research-Report.pdf

Nokian Tyres 2016, Annual Reports 2015, viewed 10 July 2016, https://dc602r66yb2n9.cloudfront.net/pub/web/attachments/publications/NokianTyres_AnnualReport_2015_lowres.pdf

NokianTyres 2016, Interim Report, 1st quarter 2016, viewed 3 July 2016, https://dc602r66yb2n9.cloudfront.net/pub/web/attachments/interim_reports/Nokian+Tyres+-Q1-2016-EN_RELEASE+final+web.pdf

Nordea 2016, Annual Reports 2015, viewed, viewed 10 July 2016, https://www.nordea.com/Images/20-104322/annual-report-2015.pdf

Pietiläinen, T. 2016, 'Entinen toimitusjohtaja myi kaikki omistuksensa Nokian Renkaista ennen rengasmanipuloinnin paljastumista', HS, viewed 3 July 2016, http://www.hs.fi/talous/a1456723410789

Pietiläinen, T. 2016, 'Nokian Renkaiden nykyjohto myi 600 000 euron optiot ennen testivilpin tulosten julkistusta', HS, viewed 3 July 2016, http://www.hs.fi/talous/a1456983422573?jako=7466dad8da31460ad438295f3e141494

Press, G. 2013, 'A Very Short History Of Big Data', Forbes, viewed 12 May 2016, http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/2/#7f05c77a1af0

Quantitative Methods in Social Scieces 2016, Validity and Reliability, viewed 11 July 2016, http://ccnmtl.columbia.edu/projects/qmss/measurement/validity_and_reliability.html

Reuters 2016,Swedish FSA investigates Nordea after Panama Papers leak, viewed 3 July 2016, http://www.reuters.com/article/us-panama-tax-nordea-idUSKCN0X10VT

Schwartzkopff, F. 2016, 'Nordea Money Laundering Link to Panama Leaks Probed by FSA',Bloomberg, viewed 3 July 2016, http://www.bloomberg.com/news/articles/2016-04-12/nordea-money-laundering-link-to-panama-leaks-probed-by-regulator

SSIX 2016, Sentiment Analysis – Business Opportunities, viewed 11 July 2016, http://ssix-project.eu/sentiment-analysis-business-opportunities/

Statista 2016, Number of social network users worldwide from 2010 to 2018 (in billions), viewed 10 Februry 2016, http://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Statista 2016, Social network advertising revenue from 2014 to 2017 (in billion U.S. dollars), viewed 10 Februry 2016, http://www.statista.com/statistics/271406/advertising-revenue-of-social-networks-worldwide/

Taboada, M., et al. 2011, 'Lexicon-based methods for sentiment analysis', Computational Linguistics, vol 37, no 2, viewed 11 July 2016, http://www.mitpressjournals.org/doi/abs/10.1162/COLI_a_00049#.V1wb6l9OJaQ

Tallink Silja Line 2016, Tallink Silja Line Company Presentation (March 2016), viewed 29 January 2016, http://www.tallink.com/documents/10192/21881052/2016-03_Tallink_Company_Presentation.pdf/56b0d0f5-47f1-47e6-b9e4-36f82360b112

Thompsons solicitors 2008, Hoisting the flag, viewed 3 July 2016, http://www.thompsons.law.co.uk/ltext/lelr-weekly-49-hoisting-flag.htm

Twitter 2016, Developer Policy, viewed 12 May 2016, https://dev.twitter.com/overview/terms/policy

Twitter 2016, Differences between Streaming and REST, viewed 11 July 2016, https://dev.twitter.com/streaming/overview

Valentine, J. 2015, 'Carmaker allegedly used software that circumvents emissions testing for certain air pollutants', EPA, viewed 3 July 2016, https://yosemite.epa.gov/opa/admpress.nsf/6424ac1caa800aab85257359003f5337/dfc8e33b5ab162b985257ec40057813b!OpenDocument

Viking Line 2015, Viking Line is launching a digital travel experience, viewed 3 July 2016, https://www.youtube.com/watch?v=lMjSB9o-xcQ

Viking Line 2016, Annual Report 2015, viewed 3 July 2016, http://www.vikingline.com/Documents/investors/annual%20reports/annual-report-2015-en.pdf

Wärtsilä 2016, Annual Reports 2015, viewed, viewed 10 July 2016, http://wartsila-reports.studio.crasman.fi/file/dl/i/MwsYMw/cnSByJl7Min8DNaPEiaQww/Wrtsil_Annual_Report_2015.pdf

Wärtsilä 2016, Interim Report Q1 2016, viewed 3 July 2016, http://www.wartsilareports.com/en-US/

YLE 2016, 150 sick: Silja Symphony ship faced with stubborn case of norovirus, viewed 3 July 2016,

http://yle.fi/uutiset/150_sick_silja_symphony_ship_faced_with_stubborn_case_of_norovirus/8926654

Yle 2016, Finnish tire firm manipulated test results, viewed 3 July 2016, http://yle.fi/uutiset/kauppalehti_finnish_tire_firm_manipulated_test_results/8702416

APPENDIX 1

Search word

1. #Fazer
2. #Finnair
3. #Fiskars
4. #Gugguu
5. #Kesko
6. #Marimekko
7. #NokianRenkaat
8. #NokiaRenkaat
9. #Nordea
10. #Outokumpu
11. #Siljaline
12. #Supercell
13. #Vikingline
14. #Wartsila
15. Destia
16. Fazer
17. Finnair
18. Finnlines
19. Fiskars
20. Gugguu
21. Ivana Helsinki
22. IvanaHelsinki
23. Kesko
24. Marimekko
25. Martela
26. Metso
27. Nokia Renkaat
28. NokianRenkaat
29. NokianTyres
30. NokiaRenkaat
31. Nordea

32. Outokumpu
33. Reima
34. Silja line
35. Siljaline
36. Stockmann
37. Supercell
38. Viking Line
39. VikingLine
40. Wartsila

APPENDIX 2

List of companies

1. Destia
2. Fazer
3. Finnair
4. Finnlines
5. Fiskars
6. Gugguu
7. Ivana Helsinki
8. Kesko
9. Marimekko
10. Martela
11. Metso
12. Nokia Renkaat
13. Nordea
14. Outokumpu
15. Reima
16. Silja line
17. Stockmann
18. Supercell
19. Viking Line
20. Wartsila

APPENDIX 3

Sentiment words list English

words.txt

APPENDIX 4

Sentiment words list Finish

words_fi.txt

APPENDIX 5

Emotion list

Emotion_list.txt

APPENDIX 6

ETL code