TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Software Engineering

Koit Summatavet

Automated Estimation of Company Reputation

Master thesis

Supervisor: Innar Liiv
Ph.D.

Tallinn 2024

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Tarkvaratehnika

Koit Summatavet

Ettevõtte maine automatiseeritud hindamine

Magistritöö

Juhendaja: Innar Liiv
Ph.D.

Tallinn 2024

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Koit Summatavet

12.05.2024

Abstract

Company reputation is a valuable commodity to companies and their stakeholders, however a universal method of measuring the reputation of companies with different types and across countries has not been agreed upon. The aim of this thesis is to create an ecosystem to measure a company's reputation by analyzing qualitative and quantitative data overcoming the limitations of existing frameworks. Based on related work the research aims to find critical attributes to measure the reputation of Estonian companies by conducting questionnaires and training models to predict future reputation. The ecosystem developed for measuring company reputation creates questionnaires to all Estonian companies, and capable of displaying the questionnaires to multiple respondents on multiple websites, resulting in data providing significant results to measure company reputation. The experimental research determined the performance of multiple models capable of predicting company reputations with high accuracy and precision.

This thesis is written in English and is 63 pages long, including 8 chapters, 2 figures and 2 tables.

Annotatsioon

Ettevõtte maine automatiseeritud hindamine

Ettevõtte maine on ettevõtete, nende klientide ja partnerite jaoks väärtuslik vara, kuid siiani ei ole kokku lepitud universaalset meetodit, kuidas ühtselt mõõta eri tüüpi ja eri riikide ettevõtete mainet. Käesoleva uurimistöö eesmärk on luua ökosüsteem ettevõtte maine mõõtmiseks, analüüsides kvalitatiivseid ja kvantitatiivseid andmeid ning ületades olemasolevate raamistike piiranguid. Valdkonnas avaldatud varasematele uurimustele tuginedes on töö eesmärk leida kriitilisi tunnuseid Eesti ettevõtete maine mõõtmiseks, viies läbi küsitlusi ja treenides mudeleid edaspidise maine hindamiseks. Ettevõtte maine mõõtmiseks välja töötatud ökosüsteemis saab luua küsimustikke kõigile Eesti ettevõtetele ja samaaegselt kuvada küsimustikke mitmetele vastajatele mitmel erineval veebilehel. Küsimustiku tulemusel kogutakse andmeid, mis annavad olulisi tulemusi ettevõtte maine mõõtmiseks. Eksperimentaalse uurimistöö kaudu treenitud mudelid on võimelised ennustama ettevõtte mainet suure täpsuse ja tõenäosusega.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 63 leheküljel, 8 peatükki, 2 joonist, 2 tabelit.

List of abbreviations and terms

CBR - Customer-based reputation

WMAC - World's Most Admired Companies

AMAC - America's Most Admired Companies

EMTAK - Estonian Classification of Economic Activities

CSV - Comma-Separated Values

SVM - Support Vector Machines

Table of contents

List of figures

List of tables

Keywords:

Statistics, analysis, programming, company reputation, financial performance, questionnaire

1 Introduction

Reputation is a valuable commodity recognized as a fundamental instrument of social order. Concerning the organization's reputation, it is linked to their identity, performance and the way people see the company. The reputation of a company has a significant impact on its value that is important to investors, it also attracts new customers and may retain the existing ones. That is why reputation is indispensable for an organization. Quoting American investor, business tycoon, philanthropist and one of the most successful investors in the world, Warren Buffet: “It can take 20 years to build a reputation, but it takes 5 minutes to lose it” (Jiang, 2020).

Organizations are commonly identified by their position in subjective quantitative rankings that either focus on a specific corporate attribute or a broader overall assessment of corporate quality (Bermiss et al. 2014). An organization’s reputation, and changes in its reputation, influence the organization’s relationships with its stakeholders (Lange et al. 2011). It is a challenge to analyze the reputation, consisting of familiarity with the organization, beliefs about what to expect from the organization in the future, and impressions about the organization’s favorability. The corporate reputation is strongly influenced also by firms’ financial performance. The social construction of corporate reputation is a result of collective sensemaking by the targeted audience and is strongly influenced by media rhetoric (Bermiss et al. 2014). Financial success has long been accepted as the primary objective of corporate existence.

However, in the last decades many organizations struggle to tell their stories, to communicate the good that they do in the marketplace, in the community, to and for the environment, and in society (Campbell and Sherman 2010; Calderón et al. 2012). Despite the interest to respect a more ethical approach to business, most firms did not put the social responsibility concept at the core of their concern, and even considered this matter as a source of additional costs (Barchiesi and La Bella 2014). The ethical aspects, sustainability and environmental issues are becoming increasingly important, and it is risky to ignore them.

Reputation is at the heart of a company’s success (Davies et al. 2003). To determine the organization’s reputation usually the opinions of stakeholders (customers, company employees etc.) are gathered for instance CBR method and WMAC framework, but this way tends to be more subjective rather than objective. Customer’s evaluation of a firm is typically based on their reactions to the firm's goods, services, communication activities, interactions with the firm and/or its representatives or constituencies (such as employees, management, or other customers) and/or known corporate activities (CBR).

The WMAC framework evaluates eight key attributes of reputation which are innovation, people management, use of corporate assets, social responsibility, quality of management, financial soundness, long-term investment, quality of products/services, and global competitiveness. These might seem to be very objective attributes, but decisions are affected by halo effect from a good financial performance (Fombrun and Shanley 1990; McGuire et al. 1990; Fryxella and Wang 1994), resulting in a more subjective method.

The first reputation ranking of US firms on a global level was Fortune’s AMAC index (Hutton, 1986). The index has later been further developed into the World’s Most Admired Companies (WMAC) index, which is annually reported by the Fortune magazine. In its original form, the AMAC comprises different criteria like financial soundness, long-term investment value, wise use of corporate assets, innovativeness, ability to attract, develop, and keep talented people, quality of products or services, quality of management, and community and environmental responsibility (Sobol and Farrelly 1988). Despite its frequent use, numerous authors criticize the AMAC index. Sobol, Farrelly and Taper (1992) note the lack of a precise construct definition and a sound theoretical grounding of the eight categories. Other authors observe that the index is subject to a strong financial halo effect (e.g., Davies, Chun and Kamins 2010; in Sarstedt et al. 2013).

Fortune magazine uses a WMAC framework to rank companies by their reputation. The problem lies with the inability to evaluate the reputation of smaller companies. Also, applying these framework costs are quite high, because of the substantial number of senior executives, board directors, and expert analysts that had to fill out the questionnaires. Lastly, this framework provided mostly subjective opinions rather than objective criteria.

Kantar Emor is a data and market research company in Estonia. The company has performed numerous studies assessing the reputation of companies and employers in Estonia. The research in 2008 included 66 companies published in year 2007 Äripäev Revenue TOP 500 companies list, where two key attributes – Estonian citizens' general attitude towards business and Estonian citizens' assessment of business strength was measured on a ten-point scale. In April of 2014 research asked 1101 participants to rank 50 most reputable employers across Estonia. The research conducted by Kantar Emor provides a large amount of data that can be used to train a prediction model for measuring company reputation, however the data is under Intellectual Property protection.

Äripäev publishes list of companies based on their revenues and financial performance in their Äripäev Revenue TOP 500 release. However, this list is behind the company's website's paywall and the data used is under their Intellectual Property protection.

Kantar Emor and Äripäev data and research covers a small group of companies which sets their methos apart from the aims of this thesis, enabling to automatically measure the reputation of all Estonian companies.

1.1 Aim and Objectives

There are numerous questionnaire platforms available. However, by evaluating different questionnaire platforms, the research focuses on developing our questionnaire to meet our criteria and cases. Websites such as Inforegister.ee, e-krediidiinfo.ee and ariregister.rik.ee provide a view for Estonian company details, where each view contains content cards relevant to that specific company enable us to integrate into the website our own developed context aware questionnaire platform.

This study aims to answer the following challenges:

- Based on related work find key attributes to measure company reputation.
- Gather reputation of Estonian companies.
- Train models to predict reputation of companies compared to one another.

Developing our methodology we follow the main objectives:

- Find the differences between the company measurement criteria and choose the most suitable based on the company data available for this research.
- Develop questionnaire application to conduct data collection.

- Choose, train and analyze models on the collected data.

This study creates an ecosystem to measure a company's reputation by analyzing qualitative and quantitative data overcoming the limitations of existing frameworks, which often rely on subjective opinions. Based on related work, the research aims to find critical attributes measuring the reputation of Estonian companies and train models predicting the reputation of companies compared to one another.

The research focuses on creating a questionnaire that allows us integration of a platform to conduct our research. Reputation is measured by choosing attributes to develop a questionnaire application to conduct data collection and train and analyze models on the collected data. The proposed method provides a more objective and scalable approach to reputation assessment.

The main argument for developing our solution of the ecosystem to conduct research in Estonia is finding relevant attributes related to the gathered company data of all Estonian companies, including the data relevant to our research (financial performance, registration code, EMTAK, active/inactive company), which makes the study more accurate and practical. To find the optimal data collection method (result questionnaire), we design a questionnaire (attribute question, companies to compare, company pair, visual solution UI/UX) to provide context and dynamic questionnaire content developing our research ecosystem to integrate the proposed questionnaires. Processing the questionnaire results for analysis, we develop and analyze models trained on questionnaire results (model selection, model training procedure, model evaluation).

The rest of the thesis is organized as follows. Chapter 2 provides an overview of related works in company reputation measurement methods, the qualitative and quantitative attributes used to assess the differences between each company, and delivery ecosystems. We determine the most relevant approaches to the companies in question and the country where the measurements were conducted.

Chapter 3 gives a brief review of the research on the data acquired by the ecosystem used to conduct the data collection and platform for the experimental research and analysis. The ecosystem was designed with accessibility to many people, and information relevant for both the visitors and the master thesis research. Explanation of the rationale for using a questionnaire, describing the architecture, unique features, benefits, potential drawbacks, scalability, and feature implementation.

Chapter 4 presents the descriptions and analyses of the datasets and technical design developed, and a detailed insight into the components and practices that contribute to the technical foundation of the reputation estimation system of the questionnaire platform, selection of appropriate software, adhering to coding best practices, managing the codebase effectively, and handling data efficiently.

Chapter 5 focuses on applications for collecting and processing all active companies in Estonia based on publicly available information, processing the relevant data and grouping companies by industry field helping to develop the company comparison method of the questionnaire.

Chapter 6 presents the results of the questionnaire distribution, data collection methods, experimental procedures, and analytical techniques used to understand the data gathered and analyzed to validate the automated estimation of company reputation.

Chapter 7 provides data analysis and results of the research and discussion.

Chapter 8 draws conclusions about the proposed models and outlines future work.

2 Theoretical Framework

2.1 Corporate reputation and stakeholders (customers, employees, suppliers, investors)

Corporate reputation has become one of the most important intangible assets for maintaining and enhancing firms' competitiveness in the global marketplace. Researchers have shown considerable interest in measuring the corporate reputation construct, resulting in a lack of consensus on valid measurement approaches (Sarstedt et al. 2013).

Corporate reputation can be broadly defined as a stakeholder's overall evaluation of a company over time. The understanding of corporate reputation is important for companies' strategic marketing activities; effective communication with different stakeholder groups is helped by understanding what people view as important components of a good reputation. Multinational corporations in particular need to examine the global consistency of their stakeholder perception (Walsh and Wiedmann 2004).

Maintaining and increasing corporate reputation has become a crucial management objective for globally operating firms. A good reputation can improve customer confidence in a company's products or advertising claims and can increase customer commitment, customer satisfaction, word-of-mouth and loyalty, and can help attract and retain talent, can limit personnel fluctuation, and increase production efficiency via lower salaries and a higher employee motivation (Sarstedt et al.2013, pp. 1).

Various fields, including economics, strategy, marketing, organization theory, sociology, communications and accounting are contributing to the burgeoning literature on corporate reputations (Fombrun and Van Riel 1997). Since corporate performance is multidimensional reflecting the unique dimensions on which individual stakeholders base their judgements of the company's performance. A corporate reputation calibrates a firm's relative standing internally with employees and externally with its other stakeholders and is therefore a collective assessment of a company's ability to provide valued outcomes to a representative group of stakeholders (Fombrun et al. 2000, pp. 242-243).

As the reputation, based on the accumulation and the sharing over the time of experiences and stories that stakeholders talk about the company (Barchiesi and La Bella 2014, pp. 162), the research and development of the internal perception and the value system of the organization become very important aspects in the field of design of company reputation ecosystems and sustainability.

2.2 Psychometric properties and intangible assets of reputation measures

In the beginning of the 21st century large firms were concerned with a wide range of stakeholders, but not very concerned with respect to the environment, society, and what the other firms in their supply chain were doing. Most attention was reserved to shareholders, implying a focus on short-term return rather than on sustainability (Barchiesi and La Bella 2014).

The companies' ability to grow and to improve continuously is also determined by its social competences, ethical responsibility, and environmental contribution. People are becoming more and more socially responsible, and hence becoming generally more sensitive to social and environmental issues (Barchiesi and La Bella 2014, pp. 164–165). However, reputation combines everything that is knowable about a firm. The philosophy is to match the external (customer) perception of the organization and what they value, to the internal (customer-facing employee) perception and their organizational values. Only when these are ‘harmonized’ can the firm be truly competitive (Davies et al. 2003).

Based on the determination process of the organization’s reputation is rather abstract and subjective than the objective evaluation process. Reputation is a comprehensive idea, which allows everything to be its explanatory condition – it is manufactured in situ and changes over time when firms are put in new contexts where their reputation is challenged (Schultz et al. 2001, pp. 38). As an empirical representation, it is a judgment of the firm made by a set of audiences on the basis of perceptions and assessments that are assembled and made available via a ranking system, which defines, assesses, and compares firms’ reputation according to certain predefined criteria. Moreover, it seems that respondents, when assessing the reputation of firms along specific criteria, project their general judgment of the firm onto new criteria for reputation. This has parallels to the ‘halo-effects’ found in other reputation studies (Brown and Perry 1994), so that the development of an increasingly complex measurement system paradoxically becomes ‘more of the same’– creating a sticky reputation (Schultz et al. 2001, pp. 25–26).

Schults et al. argue that the majority of the sticky firms are companies which seem able to adapt to the new societal expectations as to legitimate corporate behavior, with social responsibility and concern for the environment being the most obvious areas. Sticky reputation may also reflect firms’ ability to imitate new institutional demands, generating the formal structures and communicative rituals needed to obtain legitimacy (Schultz et al. 2001, pp. 37).

In order to track and improve their reputation, companies need to adequately measure their reputation and the dimensions that influence it. However, researchers point out the limitations of empirical comparison of the psychometric properties of reputation measures and used measurement models (Sarstedt et al. 2013) or measures of corporate reputation that suffer from fundamental methodological and conceptual weaknesses (Fombrun et al. 2000).

The most rapidly increasing areas of reputation are the corporate codes that have proliferated over the last decades of the 20th century. This interest in codes has grown simultaneously with the attention on corporate social responsibility and sustainable business practice in big corporations. Codes of ethics show corporate responsibility sensibility and contain a set of rules of conduct and corporate principles concerning the responsibility of a company to its stakeholders and shareholders (Calderón et al. 2012). The most difficult task is to find ways to adequately measure the empirical and experiential aspects of a company's reputation and the dimensions that influence it.

A reputation is a combination of views and impressions of many different people, not unanimously held, but general, how a company meets the expectations of all its stakeholders (Fombrun 1996; Roper and Fill 2012, pp. 5). Brands form part of the stories that help to explain life and thus have been moved from the sphere of management to the sphere of government, emphasizing their huge power and central position in people’s lives (Roper and Fill 2012, pp. 129). In addition, the rise of the service industry and the explosion of the internet have emphasized the importance of corporate rather than product branding (Roper and Fill 2012, pp. 189).

Roper and Fill (2012) point out that corporate reputation as an academic area is relatively new. They argue that research from the 1950s and 1960s looked at reputation from an external perspective, that of the customer. During the 1970s the views of staff, the internal perspective, received more attention, although both perspectives tended to be labeled as ‘image’ and the lack of clarity over an agreed lexicon has also resulted in a number of different models being produced over the years that link image and reputation. The building blocks of corporate reputation are corporate personality, corporate identity or brand, corporate image and corporate reputation (Roper and Fill 2012, pp. 34–35).

In order to build a favorable reputation, four attributes need to be developed: credibility, trustworthiness, reliability and responsibility. Based on Fombrun (1996) they emphasize the importance of the views of all stakeholders, internal and external, and not just customers (Fombrun 1996; in Roper and Fill 2012, pp. 36).

According to Roper and Fill (2012) the reputation of the company is often the key way that a customer can judge an intangible offering and Fortune 500 in the US, Management Today in the UK together with employee-based studies such as the Sunday Times’ Best Companies to Work For guide ensure that corporate reputation is high up the news agenda. They claim that a good reputation provides the organization with a competitive advantage and a company with a strong reputation will have a higher share price and more loyal customers who expect companies to behave in an ethically sound manner (Roper and Fill 2012, pp. 23).

Walsh and Wiedmann propose a model of antecedents, dimensions, and consequences of corporate reputation: emotional appeal, vision and leadership, social and environment responsibility, fairness, transparency, products and services, workplace environment, financial performance, sympathy, and perceived customer orientation (Walsh and Wiedmann 2004, pp. 310). These elements are very complex to analyze and it is crucial to understand how people interpret such tangible and intangible values from a very subjective point of view.

Internal forces such as the corporate strategy, culture and use of resources, including financial, management and employee expertise, are more controllable and can be adapted to different contexts. However, external forces are largely uncontrollable (Roper and Fill 2012, pp. 38). In order to study reputation, it is necessary to understand corporate culture as a particular area of study that explores the psychology, attitudes, beliefs and values of an organization using certain rituals and symbols is complex and something that may only be possible using qualitative research (Roper and Fill 2012, pp. 56). Psychometric properties and intangible assets of reputation measures are very important aspects to evaluate the reputation of companies.

2.3 Related work

Researchers have shown considerable interest in measuring the corporate reputation construct, resulting in a lack of consensus on valid measurement approaches. According to Lange, Lee and Dai (2011) the idea of organizational reputation is intuitive and simple in its common usage. However, it is surprisingly complex when employed and investigated in management research, as evidenced by the multiple definitions, conceptualizations, and operationalizations that have emerged across studies. Three dominant conceptualizations that reputation consists of are identified: 1) familiarity with the organization; 2) beliefs about what to expect from the organization in the future; and 3) impressions about the organization’s favorability (Lange et al. 2011, pp. 155). They argue that reputation is rooted in the organization’s historical behavior.

Sarstedt, Wilczynski and Melewar (2012) identify corporate reputation as one of the most important intangible assets playing an increasingly important role in terms of firms' propensity to influence important stakeholder groups, such as financial analysts, employees, and customers in global markets. Firstly, a good reputation can improve customer confidence in a company's products or advertising claims and can increase customer commitment, customer satisfaction, word-of-mouth and loyalty. Secondly, a good corporate reputation can help attract and retain talent, can limit personnel fluctuation, and increase production efficiency via lower salaries and a higher employee motivation (Sarstedt et al. 2012, pp. 329). Corporate reputation has become one of the most important intangible assets for maintaining and enhancing firms' competitiveness in the global marketplace. They claim that intangible assets provide more competitive advantages than product-related sources.

Corporate reputation attracts significant attention among marketing scholars. However, researchers often overlook customers' opinions specifically. Walsh and Beatty (2007) developed and tested a customer-based corporate reputation scale – the CBR scale. They define customer-based reputation (CBR) as the customer's overall evaluation of a firm based on their reactions to the firm's goods, services, communication activities, interactions with the firm and/or its representatives or constituencies (such as employees, management, or other customers) and/or known corporate activities (Walsh and Beatty 2007; Walsh et al. 2009, pp. 924).

The CBR-Short scale is well-suited to gather benchmark data regarding levels of customer-based corporate reputation as well as to conduct periodic 'checks' to measure reputation improvements.

The findings of Walsh, Beatty and Shiu (2009) indicate that a favorable (or unfavorable) corporate reputation positively (or negatively) affects critical relational outcome variables – loyalty, trust, and re-patronage intentions, and should be of concern to companies. Service firms can better identify the strengths and weaknesses of the firm's reputation in enabling management to pinpoint problems and develop solutions (Walsh et al. 2009, pp. 929). Their study emphasizes the importance of customer-based corporate reputation and highlights the need for the 'Product and Service Quality' dimension.

Park and Lee (2007) emphasize the recent popularity and pervasiveness of uncensored opinions on the internet that have created unprecedented challenges to corporate PR practitioners. Despite increased efforts, establishing and protecting positive corporate reputation is becoming increasingly difficult, because it is almost impossible to monitor all negative postings against a company available on the Internet. They examine how online postings about a company in an online news forum affect people's perception of the corporate reputation of the company in either negative or positive ways. Their tests show the relationship between the number of similar (either positive or negative) online postings about a particular company and the intensity of people's perception of the company's corporate reputation.

According to Cherchiello (2011) a detailed and systematic analysis of what the media are saying is especially important because the media shape the perceptions and expectations of all the involved actors. She proposes statistical models aimed at measuring the effective reputation of an institution. To cope with the media reputation measurement, she proposes a parametric model, and an integrated approach employing real media-based reputational data. Since reputation involves intangible assets (public opinion, perception, reliability, merit), it is not simple to define and consequently to measure and to monitor the correlated risk. The proposed models are applied to real data on Italian public companies taken from financial media corporations.

Cannon and Schwaiger (2005) address the problem by showing how company reputation could be incorporated into total enterprise simulations. They present a model for incorporating the concept of company reputation into a total enterprise simulation, and build on an empirically derived model of company reputation in which various company characteristics are linked to two underlying dimensions of reputation: sympathy and competence. According to their research, the consequence of linking simulation game decisions to reputational characteristics is twofold. Linking game decisions to reputational characteristics and reputational characteristics to specific reputational outcomes provides a powerful stimulus for learning (Cannon and Schwaiger 2005, pp. 199).

Indeed, simulation as a method is an increasingly trendy and powerful tool for both researchers and entrepreneurs. For example, Herbig and Milewicz (1995) quantify the relationship between reputation and credibility and quantitatively describe the different effects of credibility transactions on the reputation and credibility of the signaling firm. They provide a quantitative dimension to these two constructs, show their dramatically different effects on a firm’s performance, and the powerful disincentives they provide to a firm thinking about risking (milking) its hard-earned reputation for short-term gains. This is achieved by a simulation in pricing reputation.

Corporate reputation is a valuable intangible asset that needs to be managed as it influences stakeholders’ perceptions and preferences of companies as employment and investment opportunities, as community members, and as suppliers of products and services. According to Puncheva (2008) it is widely accepted that corporate reputation influences organization stakeholder interactions, but there is no theoretical framework that conceptualizes this aspect in stakeholders’ decision-making processes for establishing various forms of relationships with a firm. By adopting an interdisciplinary approach, he provides a theoretical model that explains the role corporate reputation has in the process through which stakeholders decide to establish relationships with a firm.

Although corporate reputation is a significant and relevant corporate asset, formidable measurement challenges have effectively kept this major intangible asset out of the financial statements. According to Cravens, Oliver and Ramamoorti (2003) the creation of a ‘reputation index’ that would be of broad scope and attempt to capture key dimensions and evaluate diverse organizational components including corporate strategy, financial strength and viability, organizational culture, ethics and integrity, governance processes and leadership, products/services, strategic alliances and business partnering, and innovation along with information already contained in the corporation’s annual report.

In response to the need to evaluate intangibles associated with corporate value, several methodologies have been developed that incorporate market values and book values to provide a residual measurement of intangibles. For example, Cravens, Oliver and Ramamoorti (2003) built their method and the Reputation Index on the work of Harvey and Lusch. Harvey and Lusch (1999) look at the intangibles, describing a classification and assessment index for intangible liabilities. They created separate categories for internal and external liabilities that fall into four main areas: 1) process issues; 2) human issues; 3) informational issues; 4) configuration issues (Cravens et al. 2003, pp. 209) and focus on a more comprehensive assessment of corporate reputation as a supplementary disclosure to the financial statements.

This reputation index focuses on disclosure of the positive as well as negative aspects of corporate reputation and provides an additional degree of standardization. With standardization of measures, the index would thus be more comparable across firms and industries. Much like knowledge capital, corporate reputation is an aggregate intangible asset that must be evaluated using both internal and external information (Cravens et al. 2003, pp. 204). The objective was to provide a platform from which to begin an investigation into the assessment and measurement of corporate reputation, and propose developing a composite index that relies on various scale weights that would allow standardization in the measures rather than on absolute quantitative values.

Smaiziene and Jucevicius (2009) emphasize the holistic approach to corporate reputation that leads to acknowledging the necessity for integrating human resource management into the processes of managing a company's reputation. They show that fragmentary research of reputation (individual or organizational one) can be found in psychology, sociology, economics, marketing, human resource management, impression management, business strategy, public relations, etc. According to their research, corporate reputation includes nothing else but holistic stakeholders' evaluation of a company and its processes and results on the basis of their expectations.

Holistic approach to corporate reputation enables us to compare and implement different discourse practices into a multilayered structure of the field. In their case study Sabater and Sierra (2002) used social analysis as part of the reputation system based on three dimensions of reputation: the individual dimension models (the direct interaction between two agents); social dimension (Witness Reputation, Neighborhood Reputation, System Reputation); ontological dimension. The use of social network analysis can be used in a reputation system that considers the social dimensions of the reputation.

Therefore, in addition to studies on the reputation of different companies, also the research on nonprofit organizations provides valuable empirical data. For example, Sarstedt and Schloderer (2010) adopted private sector firms' values and concepts to nonprofit organizations (NPOs), identified the dimensions of NPO reputation and developed indices to measure these components. Their model of a qualitative inquiry and a quantitative study use a large-scale sample from the German general public.

They found support for a two-dimensional measurement approach comprising an affective and cognitive component as well as four antecedent constructs (quality, performance, organizational social responsibility (OSR), and attractiveness).

They used both qualitative inquiry (sample, procedure, and interview guideline) and quantitative inquiry (Sample). Using empirical data from a large-scale survey, the model was validated and used to predict relevant criterion variables.

Caruana (1997) points out the better understanding of the role of the halo effect in reputation that can go a long way in our understanding of the resilience of reputation and its effect on other variables that include competitiveness, performance and recruitment. In his research individuals are seen to form consistent impressions of objects early in the categorization process and this is swiftly accompanied by a drastic decline in openness to new information. This feature is built on the principle of cognitive economy with response to one stimulus extended to others as a result of the "halo effect". Once established, a halo determines the perception of publics of that firm and may outlast the actual reality by quite some time. Thus, in the case of firms facing a temporary reverse of some sort, a positive halo provides a breathing space allowing the possibility of recovery. On the other hand, a firm with a poor reputation would have an uphill struggle in overcoming the initial negative halo that acts to limit the ease with which a more positive corporate reputation is established (Caruana 1997, pp. 114).

Bermiss et al. (2014) discuss the ranking of corporate reputation as a dynamic commensuration process, integrating two related streams of relevant research: 1) work on commensuration, defined as the transformation of qualitative distinctions into quantitative distinctions expressed on a common metric; and 2) work on the lifecycle of management fashions, which defines how "fashion-setters," such as the business press, influence perceptions of appropriate corporate behavior. They argue that market actors use the media to help make sense of new financial performance indicators, particularly when there is no general consensus as to which paradigm is the most accurate. Their research highlights the social constructionist perspective of the change of beliefs about the legitimate indicators of firm performance shaping rankings and influenced by prominent market intermediaries (e.g., the business press).

Campbell and Sherman (2010) explore the relationships between and among lists of top performers – Boston College's Corporate Social Responsibility Index, Fortune's World's Most Admired Companies, Harris Interactive's Reputation Quotient for the Most Visible Companies, and Newsweek's Green Rankings of America's 500 largest corporations. Their research enables them to explore whether in spite of differing methodologies and criteria for rankings, there is a commonality of inclusion in lists. They focus on people's awareness about a company, what value the company has created in the past, is creating currently, and will create in the future. In addition, they emphasize much broader scope of criterias of the Most Admired Companies lists that follows nine attributes: ability to attract and retain talented people; quality of management; social responsibility to the community and the environment; innovativeness; quality of products or services; wise use of corporate assets; financial soundness; long-term investment value; and effectiveness in doing business globally (Campbell and Sherman 2010, pp. 48). They claim that people look to different sources to form an opinion of how a company is fulfilling its obligations to other stakeholders.

Barchiesi and La Bella (2014) examine the value statements of the companies that have kept their place within the top 50 of the world's most admired companies' rankings, made by the authoritative business magazine Fortune, from 2009 to 2013. They identified five independent core value orientations: to customer/user, to employees, to economic and financial growth, to excellence, and to social responsibility. Their analysis highlights the relation between high revenues, global admiration, and social-responsibility-oriented core values. Their research explains how social responsibility plays a key role as an underlying guide to many companies that achieved recognized excellence in their fields of activity (Barchiesi and La Bella 2014, pp. 160). In their study, they investigate how deeply social responsibility is rooted in the core values of the world's most admired companies to propose a taxonomy of organizational core value orientations. Their analysis confirms the strict relation between high revenues, global admiration, and social responsibility. They suggest reputation is sticky over time despite shifting ranking criteria and fragile statistical methods.

Schultz et al. (2001) analyze the mechanics used to construct the studies of reputation ranking systems, based on a study of a Danish ranking system, similar to the Fortune system. They show that the ongoing development of an increasingly complex measurement system paradoxically becomes ‘more of the same' and thus creates a sticky reputation for large and visible companies. According to their research, sticky reputation is durable and tends to reproduce itself over time. The reputation is about perception and interpretation of various cues and the audiences cannot see these specific improvements and the linkages between people and scoring remain problematic as we discover empirical relatedness between criteria. They point out that the longitudinal constitution of reputation is complex – it includes attention to an unknowable set of possible hunches and guesses, which depends not only on the set of respondents, but also on the set of questions, the set of methods used, and the set of statistical procedures preferred. Thus, it may be that the ultimate explanation for reputational stickiness cannot be found. It may be an inexplicable, dynamic, and self-referential system of reputation building (Schultz et al. 2001, pp. 38).

Sarstedt et al. (2013) evaluate the degree to which the reputation measurement approaches correlate with relevant outcomes such as loyalty, word-of-mouth, trust, and customers’ behavioral intentions. By examining the measures’ psychometric properties (loyalty, word-of-mouth, trust, and customers’ behavioral intentions), both theoretically and empirically, the study provides guidance for their reasonable application in business research and practice. According to their research the AMAC index is subject to a strong financial halo effect and shows that the index mainly captures a company’s past financial success rather than its corporate reputation in terms of the stakeholders’ overall evaluation (Sarstedt et al. 2013, pp. 3). They established separate path models for each combination of reputation measurement approach and criterion variable to assess the reputation measures’ criterion validities. The assessment of the criterion validities draws on 35 different model set-ups: Overall reputation, Satisfaction, Loyalty, Trust, Word-of-mouth, Commitment, Customer citizenship behavior: helping the company, Customer citizenship behavior: helping other customers.

In addition, the most admired companies of the world have codes settled on social, global and environmental aspects of corporate responsibility, paying attention to external stakeholders beyond just shareholders and internal stakeholders. For example, Calderón et al. (2012) argue that the philosophy of corporate social responsibility (CSR) is scarcely present in the codes of the most reputable companies to define the moral compass that must be used to measure reality. They examined the Most Admired Companies of the World, ranked by Fortune magazine in 2009 from two aspects: 1) whether their codes of ethics exhibit greater emphasis on social responsibility and strong implementation processes, and 2) whether they could be considered codes of the third generation. They claim that the content of the code is always a reflection of what companies hold desirable and this defines the moral compass that must be used to measure reality.

Fombrun et al. (2000) propose 'the reputation quotient' (RQ) – an instrument to measure corporate reputations and establish its empirical validity and reliability through focus groups and pilot studies. They claim that the reputation quotient is a robust measure of corporate reputations that considerably improves the state of the art in reputation measurement. Different ratings (Fortune, KLD) routinely acknowledge many limitations to the data, most also recognise the difficulty inherent in developing a valid database of corporate reputational ratings (Fombrun et al. 2000, pp. 242–244).

Fortune AMAC, released in 1984 based on data collected during the autumn of 1983, provides the longest continuing measure of corporate reputation and is the data backbone for most empirical academic work on reputation (Fombrun et al. 2000, pp. 243). Participants are asked to assess all firms in their industry on eight criteria: quality of management; innovativeness; long-term investment value: financial soundness; ability to attract, develop and keep talented people; responsibility to the community and the environment; and wise use of corporate assets. Manager Magazine (MM) ranks the companies on five attributes: management quality; innovativeness; communication ability; environmental orientation and financial stability (Fombrun et al. 2000, pp. 244). Management Today (MT) measure Britain's ten largest public companies in 26 sectors on nine criteria: quality of management; financial soundness; ability to attract, develop and retain top talent; quality of good and services; value as a long-term investment; capacity to innovate; quality of marketing; community and environmental responsibility; and use of corporate assets. In addition, their study includes Asian Business (AB), Far Eastern Economic Review (FEER), Financial Times (FT), Industry Week (IW), Fortune GMAC.

Items from existing reputation surveys: ability to attract, develop and retain top talent; ability to cope with changing economic environment; being honest and ethical; best practices – markets; business leadership; companies that other try to emulate; contribution to local economy; financial soundness; globalisation of business; innovativeness; innovativeness in responding to customers; long-term investment value; potential future profit; long-term financial vision; maximising customer satisfaction and loyalty; overall admiration; overall awareness of company; overall leadership; potential growth; quality of management; quality of marketing; quality of products and services; ranked attributes in order of importance; robust and human corporate culture; social responsibility (society, environment, community); strong and consistent profit performance; strong and well thought out strategy; use of corporate assets (Fombrun et al. 2000, pp. 246).

Fombrun and van Riel (2003) claim that reputations reflect how companies are perceived across a broad spectrum of stakeholders and a good reputation matters because it is a key source of distinctiveness that produces support for the company and differentiates it from rivals. According to Heugens (2004) corporate reputation that is so distinctive to Fombrun and van Riel's other work suggests a pluralistic answer: reputations attract better employees by making jobs more prestigious, stimulate customers to make repeat purchases, lower the cost of capital by making the firm a more credible investment object, soothe journalists by means of an overall positive halo, and command more favorable coverage by financial analysts (Heugens 2004, pp. 391).

However, Fombrun and van Riel (2003) explain how the measure of reputation that is being used in different studies is composed of six dimensions, referring to a firm's emotional appeal, products and services, workplace environment, financial performance, vision and leadership, and social responsibility (F&F 58–60). Fombrun and van Riel's central prescriptive offering, the StellarRep model (F&F 85–101). The model consists of a five-pointed star with five organizational qualities at the tips: visibility, distinctiveness, authenticity, transparency, and consistency. According to the authors, companies with high Reputation Quotient scores appear to be substantially different from lesser rivals on these five dimensions (F&F 86). Fame & Fortune's Managerial Prescriptions – Visibility, Distinctiveness, Authenticity, Transparency, Consistency. Very specific reputational problems, respectively: (a) how to maintain a stellar reputation?, (b) how to create a more visible reputation?, and (c) how to restore a damaged reputation?

Walsh and Wiedmann (2004) evaluate the cross-cultural applicability of the Reputation Quotient (RQ), an instrument to measure corporate reputation that is not biased towards one stakeholder group (e.g., financial analysts). They focus on the lack of previous comprehensive reputation research in Germany, together with the need to test the generalizability of the RQ in different countries, prompted an investigation of German stakeholders. They argue that the original US RQ could not be confirmed completely, but qualitative support was found for an extended ten-dimensional RQ, including four new German dimensions, namely: fairness, sympathy, transparency and perceived customer orientation.

Roper and Fill (2012) highlight two specific tools of reputation measurement: Fortune's 'Most Admired Companies' and 'Britain's Most Admired Companies' (Roper & Fill 2012: 82–89). Measuring tangible and intangible facets of corporate reputation, Harris-Fombrun Reputation Quotient – the RQ has been used to measure reputation in 26 countries. This system measures reputation by asking respondents to measure companies on 20 items which group around six dimensions: emotional appeal, products and services, vision and leadership, workplace environment, financial performance, social responsibility (Roper and Fill 2012, pp. 91).

Serbanica and Popescu (2009) implement a well-known model to evaluate corporate reputation – Fortune World's Most Admired Companies and Institute for Reputation Model (RepTrack® Pulse) to evaluate Romanian companies. In a developing country, with little experience in an open market, some of the so-called universally valid criteria are very new and unknown to respondents (e.g., corporate governance), while others are irrelevant (e.g., capacity to act globally).

Barchiesi and La Bella (2014) examine the value statements of the companies that have kept their place within the top 50 of the world's most admired companies' rankings, made by the authoritative business magazine Fortune, from 2009 to 2013. They identified five independent core value orientations: to customer/user, to employees, to economic and financial growth, to excellence, and to social responsibility. Their analysis highlights the relation between high revenues, global admiration, and social-responsibility-oriented core values. More recently, this concept has evolved into a broader field of studies and practices labeled as corporate social responsibility (Garriga and Melè, 2004; Costa and Menichini, 2013; Global Reporting Initiative, 2014), which has at its center three basic principles: sustainability, accountability and transparency.

Schwaiger (2004) points out that corporate reputation may cause sustainable profits, increasing competition in a globalized economy promotes the identification of drivers of sustainable competitive advantages in the field of intangible assets. Thus, evaluating corporate reputation not only appraises subjective perceptions of a company's attributes (such as "successful company", "high quality products" and so on) but also allows an intrinsic disposition towards these attributes (in the sense of "this company is not that successful, but I like it anyway", or vice versa) this research is treating corporate reputation as a two-dimensional construct. Because corporate reputation is based on perceptions far more than on concrete knowledge, managing corporate reputation is not only, but primarily, a task of corporate communications.

According to Schwaiger: Until 1997, Fortune's AMAC (America's Most Admired Companies) was the only reputation ranking available on a global level, but it was restricted to US firms. Only in 1997 did Fortune publish the results of a survey on the Global 500, divided into 24 industries and 13 countries and named GMAC (Global Most Admired Companies). In fact, highly regarded companies were not often global players, but something like multinational operating conglomerates. Since 1983, about 8,000 people have been interviewed on a regular basis on behalf of the Fortune magazine via phone and mail. The rate of return was about 50% in 1985. Survey: Respondents were asked to name the leading firms in their economic sector (rating is allowed only within the industry), and asked “How would you rate these companies on each of the following attributes”: innovativeness, quality of management, long-term investment value, community, and environmental responsibility, ability to attract, develop, and keep talented people, quality of products or services, financial soundness, and use of corporate assets. The “Overall Reputation Score (ORS)” is the arithmetic mean of the attribute's respondents provided on eight 11-point scales.

3 Methodology

3.1 Methods and empirical data

We may group the methodological contributions on measuring reputation in two categories: 1) qualitative methods; and 2) quantitative methods. However, recent studies implicate the use of different research methods and multidisciplinary approach of theory and practice in the field.

For instance, Tischer and Hildebrandt (2011) focus on a comprehensive theoretical background, why reputation has to affect financial performance. In their opinion, investors gain new information from the published rankings (increase or decrease in reputation) to adjust share prices. They developed the theoretical basis for the research by relating corporate reputation and shareholder value. They define corporate reputation as a relatively stable, aggregated and indirectly suggestible perception within multiple stakeholder groups based on a company’s past actions and future prospects in comparison to some reference. Their analysis demonstrates two distinct things based on the data used. On the one hand, publications of reputation rankings have an impact on shareholder value. On the other hand, neither good nor bad reputation scores in a ranking, nor their changes are solely appropriate to generate excess returns in the long run.

Cees van Riel (2005) introduces RepTrak® instrument developed from qualitative research conducted in 7 countries. RepTrak® is built on 22 attributes organized around 7 dimensions, including dimension scores that are made up by averages of 3-4 attributes each. The attribute weights differ from country to country and are determined based on multivariate analyses conducted from RI’s national reputation studies.

Ponzi, Fombrun and Gardberg (2011) introduced an improved RepTrak® Pulse instrument with the intention to capture corporate reputation’s emotional component. The measurement relies on three attributes that are designed to elicit a company's emotional appeal to participants including a fourth question for rating a company's overall reputation. Each item was measured with a 7-point scale, where 1 equals ‘Strongly Disagree’ and 7 equals ‘Strongly Disagree’.

Watson and Kitchen (2010) demonstrate that reputation is at the heart of all organizations, irrespective of stakeholders’ perspectives as to whether these organizations are good or bad. In the corporate world, reputation is seen as a major element of an organization’s provenance alongside and included in financial performance and innovation. The reputation relates to leadership, management and organizational operations; the quality of products and services; and – crucially – relationships with stakeholders. It is also connected to communication activities and feedback mechanisms. It is also linked to the organization’s identity, performance and the way others respond to its behavior.

The aim of Gotsi and Wilson (2001) is to clearly define the concept of corporate reputation and identify its relationship with corporate image. They attempted to make an advance in the understanding of the concept of corporate reputation by specifying its relationship with the construct of corporate image and by adopting a clear definition for future reference. They argue that the corporate reputations that a firm has with its stakeholders must rather be regarded as dynamic constructs, which influence and are influenced by all the ways in which a company projects its images: its behavior, communication and symbolism.

Coyne (2010) demonstrates that the interaction of personal and unit reputation can influence how engaged employees are in their work role as well as the amount that employees identify with a group. In addition, the interaction of unit reputation with an individual's personal reputation can have consequences on in-role performance behaviors and employee mobility. The development and maintenance of unit reputation can be influential in affecting the effort, satisfaction, and extra-role behaviors of employees.

Lange, Lee and Dai (2011) emphasize multidimensionality of approaches to defining organizational reputation. They describe organizational reputation as a multidimensional construct of 3 conceptualizations of organizational reputation: 1) being known for something; 2) being known for something and 3) generalized favorability. According to their research the management scholars have not addressed the relationship between organizational reputation and the status or reputation of the organization's affiliates (Lange et al. 2011, pp. 180).

In order to track and improve their reputation, companies need to adequately measure their reputation and the dimensions that influence it. Sarstedt, Wilczynski and Melewar (2012) claim that measuring reputation is of great importance for researchers who seek to examine its role as an antecedent, criterion, or moderating variable in different contexts. They emphasize the lack of consensus on valid measurement approaches as the biggest barrier to an effective reputation management, and there is no empirical comparison of the psychometric properties of reputation measures. First, they critically review several measurement approaches, considering recent measurement theory research and current findings from the reputation literature. Second, using an empirical survey in the German mobile phone sector, they compare the measures in terms of convergent validity, which is a fundamental criterion of any measurement instrument. They adopt the criterion validity and evaluate the degree to which the reputation measurement approaches correlate with relevant outcomes such as loyalty, word-of-mouth, trust, and customers' behavioral intentions. Their research shows that examining these associations is crucial because customer relationships (as indicated by, e.g., satisfied and loyal customers) positively affect a firm's future net cash flows and marketplace performance (Ibid., pp. 330).

The method of their research included three aspects: 1) Research design: customer satisfaction, customer loyalty, trust, word-of-mouth, commitment, and two types of customer citizenship behaviors; 2) Model estimation and comparison: overall reputation, satisfaction, loyalty, trust, word-of-mouth, commitment, customer citizenship behavior (helping the company), and customer citizenship behavior (helping other customers); 3) Data and sample characteristics: a total of 2200 randomly chosen students from all fields at a major German university were contacted via email and invited to participate in a ‘‘study on mobile phone providers.” A total of 306 respondents completed the survey fully, yielding a response rate of 13.91%. However, in looking for sources of unique and imperfectly imitable competitive advantage, both researchers and practitioners have turned their focus towards the examination of intangible assets. As a result of the study, measuring reputation formatively provides managers with important means for taking actions, not only to increase reputation itself but also to effectively manage the consequences of reputation such as customer loyalty, trust or word-of-mouth (Ibid., pp. 337).

Serbanica and Popescu (2009) combine these two methods. Their study shows that in a developing country, with little experience on an open market, some of the so-called universally valid criteria are very new and unknown to respondents (e.g., corporate governance), while others are irrelevant (e.g., capacity to act globally). There are also some more relevant evaluation criteria for Romanian companies: seriousness, degree of modernization, good salaries or compliance with European standards. At the same time, the study parallels the qualitative study results with statistical data describing the country profile to demonstrate that perception companies have roots in the socio-economic and cultural context. There is strong evidence that a good reputation leads to the clients' and employees' loyalty, to investors' retention or to a better social and financial performance. This study demonstrates that the perception over corporate reputation strongly relies on the national culture, the state of the socio-economic development and on the market condition.

Jin, Park and Kim (2008) focus on cross-cultural examination of the relationships among firm reputation, e-satisfaction, e-trust, and e-loyalty. They empirically compared the impact of firm reputation on consumers' evaluation of e-tailers' market response outcomes (satisfaction, trust, and loyalty) in two cultures – the USA (individualism, low uncertainty avoidance, low context, and high-trust society) and South Korea (collectivism, high uncertainty, high context, and low-trust society). They found that the firm reputation-trust-loyalty link is the same across cultures. However, the firm reputation-satisfaction-loyalty link is stronger in Korea than in the USA. While it may be premature to conclude the link is stronger in all Asian markets, international managers should carefully consider this finding when establishing operations in Asian markets.

Juliusson and Eriksson (2011) emphasize that in today's corporate world, where the marketplace demands the highest of standards on companies and their products, it is of detrimental nature that consumers possess trust for the brands that define their lives (incl. IKEA, SAAB, Automotive and SJ) which are continuously under intense media pressure and were therefore chosen to be the focal cases for the study. The comparative analysis of these cases proved the central importance of the product, the importance of the consumers' relationship to the brand and the importance of sound corporative governance and management in the creation of brand trust among consumers today.

Cherchiello (2011) lists among the qualitative approaches four indicators, often employed in actual corporate contexts: 1) Reputation Quotient; 2) Reputation Index; 3) Fortune's Most Admired Companies; and 4) RepTrak®. The study combines six conceptual dimensions: emotional appeal, products and services, vision and leadership, workplace environment, social and environmental responsibility, and financial performance. She emphasizes the quantitative approaches that tend to overcome the weaknesses of the qualitative methodologies. The most recognized methodologies are: 1) Intellectual Capital approach; 2) Accounting approach; and 3) Marketing approach. The Intellectual Capital approach is based on the appropriate estimation of 5 dimensions: trademark, service marks, copyrights, authorizations and exclusive rights. The Accounting approach is based on the evaluation and analysis of intangible assets, thus it is necessary to introduce criteria for fair value assessment. Finally, the Marketing approach suggests measuring the brand of a company. The more objective method considers the royalties a company can gather by conferring its brand (Cherchiello 2011, pp. 60–61). In order to evaluate the corporate reputation, she proposes a parametric statistical model whose estimation allows not only to describe and rank reputation, but also to predict and, therefore to prevent, reputational risks.

Walsh, Beatty and Shiu (2009) identify dimensions of customer-based corporate reputation (CBR); they develop scales to measure these dimensions. Researchers in the present study use the Walsh and Beatty CBR scale in the UK and Germany across contexts to study the cross-cultural validity of the measure of customer-based corporate reputation. This study assesses an abbreviated version of the CBR scale (with 15 items). The CBR Short scale has equally good dimensional properties as the original scale. They propose the psychometric data for the 28-item and 15-item CBR scales that constitute five subscales, tapping the five domains: 1) Customer Orientation; 2) Good Employer; 3) Reliable and Financially Strong Company; 4) Product and Service Quality; and 5) Social and Environmental Responsibility (Walsh et al. 2009, pp. 925–926).

Cravens, Oliver and Ramamoorti (2003) describe how an index and the preliminary components of the corporate reputation index should be developed. After illustrating the various components of the index, a methodology is provided as to the mechanics used to arrive at a summary evaluation metric. To transform the qualitative measures of their research into a form more suited to computing a reputation index suggest a nine-point scale to assess the magnitude of the measure, and list the elements that would be measured as the Components of the Reputation Index: products, employees, external relationships (non-customer), innovation and value creation, financial strengths and viability, strategy, culture, and intangible liabilities. They claim that with a reputation audit, management can address potential concerns from the public regarding the motives for disclosure and completeness of information supplemental to the financial statements (Cravens, Oliver and Ramamoorti, 2003, pp. 206–207).

Public perception on the corporate reputation of a particular company can be easily influenced by opinions on the Internet. Park and Lee (2007) in their study test the effects of online discussions about a company in an online news forum on people's perception of the corporate reputation of the company. The study is a 2 (tone of comments: negative versus positive) by 2 (number of comments: one versus five) factorial design experiment (n = 80). They found significant interaction effects between the two factors (tone versus number of comments) with regard to people's perception of the company's social responsiveness and employee treatment.

Williams and Barrett (2000) examined the influence of corporate giving programs on the link between certain categories of corporate crime and corporate reputation. They emphasized how Fombrun and Shanley (1990) analyzed the correlations among the eight attributes, and summarized their results with a rank order of variables that affect reputation and reflect the eight categories, among the variables included were financial performance, risk, size, institutional ownership, dividend yield, media exposure, diversification (strategy), and social responsibility. Williams and Barrett included these same variables in their study as the control and independent variables and added an additional independent variable for criminal activity in order to extend the prior studies. Based on two previous research (Brown and Perry 1994; Fombrun and Shanley 1990) they confirmed that firms' reputation scores are heavily influenced by the firms' financial performance.

Brown and Perry state that firms' financial performance may blur distinctions among dimensions of reputation due to a strong overall impression, or "halo" effect (Brown and Perry 1994, pp. 1349).

The presence of a "halo" effect in which firms' financial performance may heavily influence the firms' reputation scores and blur distinctions among the categories of reputation was considered, and how a dual approach was taken to ensure that the presence of a halo effect would not unduly bias the findings of the present study. Given the presence of a halo effect, they initially utilized the two-stage regression technique proposed by Brown and Perry to obtain "halo-free" data (Williams and Barrett 2000, pp. 344–346).

They found that the two-stage regression technique generated similar results to those produced by simply including the performance measure as a control variable in the regression equation. They opted for the more traditional approach of adding control variables for firm performance to their regression model and utilized OLS regression with the dependent variable (REPSCORE) regressed on the two independent variables and the seven control variables. A forced order of entry technique was utilized, with the control variables entered first in the equation. The entire sample (184 firms) was used to test H1 and H2. A hierarchical, moderated regression analysis was used to test Hypothesis 3. In the base model, the dependent variable (REPSCORE) was regressed on the two independent variables and the seven control variables. Next, we added to the base model the interaction of CITATIONS with GIVING. Since H3 examines the moderating influence of GIVING on the illegal activity-reputation relationship, we decided to use the abbreviated sample (168 firms) containing only those firms cited for criminal violations. The increment in R2 resulting from the inclusion of the interaction term in the regression equation was used to assess the significance of the interaction (Williams and Barrett 2000, pp. 345–346).

3.2 Questionnaire Design and Implementation

3.2.1 Rationale for Using a Questionnaire

This thesis follows the works by Fombrun et al. developed RepTrak® method and the Fortune magazine's published AMAC, for their main methodology of gathering data using questionnaires. These methodologies were taken as inspiration and analyzed, which led to the choice of questionnaire as the method to gather qualitative data for the thesis.

The method we use to find a company's reputation is gathering comparison data between two similar companies and finding where in terms of popularity or favorability a certain company will fit. The basis is that if we have many answers between two similar companies in the same field, then based on the number of decisions, we predict one company to be more favorable than the other. We do not compare a company with all the companies in the same industry field, but rather a set of companies from that field. This reduces the number of answers needed to make a statistical prediction of where a company falls regarding those similar companies.

A valuable piece of information about all the companies is their financial status. Many authors who have researched company reputation have checked if and how much of an impact does the financial performance of a company influences how people perceive and thus rate the company. Thus, one of the results of this thesis is that we will see whether financial performance has had any impact on the decision – whether people would rather choose the company that has a better financial performance.

3.2.2 Comparison with Other Methods Using Questionnaires

The initial idea was to use a popular questionnaire option – Google Forms. It appeared to be a good solution because it had a ready platform to create your own questions and receive the answers by email or check them in a list. The first apparent problem was how to customize the form questions to be about the specific company that the user was visiting. Google Forms has specific and not flexible form fields and thus this was not ideal for creating a form that has changeable company name and main question fields. Another problem was that the visitor had to write the other company's name by hand, which meant extra manual work for the visitor. The probability of the visitor just dismissing the questionnaire was high and therefore, using the Google Forms platform for the questionnaire was abandoned.

The same customizability issue became apparent with other questionnaires because most of them are designed to have a form that doesn't change after creation.

The next solution was to create the questionnaire ourselves to make the choice straightforward and more pleasing for the visitor, as well as to create a more flexible design and question possibilities for us. This way, we can make a questionnaire that fits the website style and has questions more relevant to the current situation.

3.2.3 Evaluation of Different Questionnaire Platforms

There are numerous questionnaire platforms available, each unique with their pros and cons. Out of the most popular and accessible options we have analyzed the following questionnaire platforms: Google Forms, SurveyMonkey, Typeform and SurveyDaddy.

Google Forms offers a free and approachable solution to create static questionnaires with the limiting factors being lack of flexibility, customizability and losing the possibility of changes once the form is created.

SurveyMonkey offers a free and a paid tier option from which the free tier is applicable for simple questionnaire cases, however, lacks flexibility, customizability and re-editing. In addition, the free tier is limited to 10 questions and 100 responses per survey.

Typeform and SurveyDaddy both have the same options and limitations as SurveyMonkey has.

Questionnaire forms that have fields that need to be filled in manually by the visitor will have the following issues: time consuming, prone to errors, fatigue and impatience. This significantly increases friction for the visitor which leads to a higher rate of dismissal.

Due to these limitations, we can conclude that these platforms are not applicable to our use case.

3.2.4 Unique Features of Our Questionnaire

To improve the user experience compared to manually filled questionnaires, we implemented a tool to display a comparative decision questionnaire with a qualitative question and two companies to choose from. The company can be chosen by pressing on the visual button component which will record the choice. In addition, we display the results of all the questionnaire decision by other visitors regarding that specific company pair. Finally, we ask the user to compare another company pair. These unique implementations increase user retention.

The main questions can be based on the most significant qualitative questions. The customizability of the main question adds more value to the questionnaire because we can cover additional characteristics of the companies. These characteristics can be qualitative attributes listed by Fombrun and Shanley (1990): financial performance, risk, size, institutional ownership, dividend yield, media exposure, diversification (strategy), and social responsibility. However, for our thesis we chose the question used by Fombrun et al. developed RepTrak® method: company's overall reputation.

Another feature implemented is displaying a unique pair only once to that specific visitor. This way, we will always ask the visitor to rate a statistically relevant pair and eliminate the same comparison to appear more than once.

Visitors can decline to answer and hide the questionnaire completely. The questionnaire won't be displayed to the user for an hour on every page visit. This improves user experience and decreases friction.

3.2.5 Benefits of Our Approach

One of the most significant benefits of the questionnaire we created is that the form does not need any code change after it's deployed if we want to add new questions or companies or update the texts. This functionality is enabled by storing all the data on a database. Thus, updating the data is quick and will not cause downtime.

Also, if any changes are needed in the code or visual representation, we can implement them ourselves. We govern and maintain the codebase.

We store the created and collected data, and therefore we handle the data storing regulations and policies, e.g. The European Union's General Data Protection Regulation (GDPR). This means we can adhere to data collection policy as soon as needed. Also, we have access to the data at any time. We do not share stored data with any third party. We can govern what data is stored and only people with authorization have access to the stored data.

Regarding the type of content, the website provides to their customers, our questionnaire not only fits a website such as Inforegister's company lookup service, but any site that might mention or contain information about a company. The website can be any news or a stocks site, it can also be a separate browser application that handles the detection of companies on the webpage; the questionnaire can be a modal in a mobile application - any site that can display a modal or a content card.

When the data is stored on our services then we can create APIs for data gathering, enabling customized API endpoints returning the data in the format required by any of our use case.

With the questionnaire platforms mentioned in paragraph 3.2.3, the data storing regulations and policies are managed by the platform and applications to data removal need to be applied through their provided services. Also, we need to inform the questionnaire participants about this. This adds additional steps for data removal, and we cannot guarantee data being sold to a third party. Furthermore, data breaches to the platforms have happened and this sets our data at risk. The platforms do not have customizable APIs for data gathering, which adds additional steps to gather and customize that data depending on our use case.

3.2.6 Potential Drawbacks

One drawback of creating our own questionnaire platform is maintenance, and it is our responsibility to update and fix the questionnaire in case of issues.

The questionnaire is created by us and thus can be configured to be integrated into any website. However, the biggest issues can be browser or website incompatibility. A website compatibility issue is present when a website has been built using website builders without customization and lack of control over the website's codebase. Also, we need to provide a fallback method in case of browser incompatibility. In the case of maintenance there can be server downtime that can lead to issues with the front-end UI, which heavily relies on the content that will be displayed.

To limit the chance of answering the same questionnaire pair more than once we assign a unique cookie to each user. However, cookies can be blocked or removed which breaks this functionality. Due to this, it is difficult to ensure that every questionnaire answer is unique meaning answered only once by a user.

3.2.7 Scalability and Feature Implementation

The questionnaire is a separate service from the website and is built on a serverless architecture. This means that the questionnaire can be integrated into any website that can display a modal or a content card. The questionnaire is built to be flexible and can be customized to fit the website's style and content. Content targeting is crucial for the questionnaire to be presented at the right time, on the right page and to the right person. The website that the questionnaire will be integrated into needs to be oriented for people interested in specific companies. The page should be able to display information about a specific company and the questionnaire.

We have implemented our database for storing and fetching data for the questionnaire. Thus, we can add, remove, or update any content based on our needs. This includes adding new questionnaire questions, companies, and updating the questionnaire's visual representation. We can also add new features to the questionnaire, such as the ability to hide the questionnaire for a certain time or the ability to decline to answer the questionnaire.

Due to the questionnaire being a separate service from the website and the serverless architecture, if the traffic to the questionnaire increases, we can scale the service horizontally and add more servers to handle the traffic.

Using the free tier database provided by mLab we had a set size restriction of the database at 500MB. If the database size exceeds 500MB, we can upgrade the database to a bigger size, however for an added cost, we can migrate the database to another service that has a bigger database size limit at better pricing.

3.2.8 Questions Included in the Questionnaire

The questionnaire platform developed in this thesis uses the question from Fombrun et al. developed RepTrak® method: company's overall reputation. More specifically we ask which one out of two companies related to each other by industry has a higher reputation? One of the companies is the company the user is visiting, and the other is randomly selected from the set of pairs created in relation to that specific company based on its industry field.

As previously mentioned, our questionnaire platform has the benefit of adding additional questions whenever needed. Because the questionnaire's front-end implementation is embedded on the website and is not built into it, it then makes it possible to configure the questionnaire based on our use case and not impact the website where the questionnaire is displayed.

3.3 Questionnaire Delivery

3.3.1 Company Pair Creation Service

The service for creating the pairs has been developed separately from the front-end and the back-end codebase of the questionnaire. Its only functionality is to create pairs from a data document, containing a list of companies and additional fields, and then record the processed documents in the database.

To create the list of companies, public documents published by the Estonian tax office are used to find the company register codes, company name, EMTAK code and the revenue. Next, we list the data we extract from the following fields provided in the documents:

From revenue documents: registration code; name; type; registered in the VAT register; EMTAK sector of activity indicated by an alphabetic code in the EMTAK structure; county; national taxes; labor taxes and payments; turnover; employees.

From business registry: name, business registration code, KMKR number, business status, business status in text, business address, location in business address, location address code, location address code in text, index in business address, ads adr id, ads adr oid, ads normalized full address, information system link.

Once all the data is read in, we need to check for companies that are missing any of the required data after which those will be filtered out. In our data we have enough information (all have a name, reg. code, EMTAK and gross amount) for about 89 000 companies.

To create the company pairs we need to divide all the companies into groups. We can use EMTAK codes to create the first categorization. This way we have companies from similar industries to each other in the same group. By having companies related to each other by industry field, we can improve the comparison and get more accurate results. Comparing companies from different fields will not give a meaningful result, for example comparing an oil company to a bank will not provide us with a statistically significant yield.

Once all the companies are divided into groups based on their EMTAK codes, we will create another set of groups where we divide the companies in the same EMTAK field based on their financial performance (gross amount). We will produce three groups where the first group will contain one third of the highest gross amount companies of the group, the second group will have a middle one third of the companies and finally the third group will have the companies with the lowest gross amount.

Depending on the number of companies in each of the groups we separate the groups further. In the first group we have EMTAK code groups that have less than or exactly 5 companies, then to the second group we place the EMTAK code groups that have less than 10 companies excluding the companies in the first group, and finally we have the EMTAK code groups that have more than 10 companies. This separation is needed for different methods of pair creation: one where all the companies in the same group are used as the second pair; another where the first three, middle three and last three companies are chosen as the second pair. A pair made by the same company will be ignored.

Once the groups have been developed, we will follow the previously mentioned separation and create a pair for each company in the list of 89 000 companies. In the created pair data, we will have the following fields: company 1 name, company 2 name, company 1 registration code, company 2 registration code, company 1 gross amount, company 2 gross amount. The service will then create a comma separated .csv document with all the pair data.

The reasoning behind having different pair creation is to create enough pairs for the comparison and not to create too many, otherwise, the results would not create a measurable and repeatable event and lead to saturation. Meaning, if the measurement is accurate then we can predict or assume the reputation for other companies belonging in the same EMTAK field, thus giving us a good model of how company reputation works. In addition, we decrease the number of pairs the user must answer and give a better idea of where the company suits. Whether the company fits the first, second, or third group based on their financial performance. Another finding is whether financial performance has any impact on the user answering the questionnaire. If most of the answers fall into these groups, then either the impact of financial performance is high (or present; answers are subjective and based on their financial performance) or the company reputation actually is high and has no impact from financial performance.

3.3.2 Providing the Company Pair

The questionnaire needs the following from the website where it is embedded: the currently relevant company registration code. Questionnaire uses this to select the pairs created for that company. As mentioned in section 3.3.1 there is a certain set of pairs created for 89 000 companies and if the current company being visited does not have any pairs for comparison, then the questionnaire won't be displayed for the visitor. This can happen when the company does not have any other companies in the same EMTAK field. In case the visited company has any pairs available it will display a pair randomly selected from the possible pairs for that company.

There are currently many websites that specialize in providing information about companies in Estonia and therefore any one of them could be the platform to display the questionnaire. One such pages is Inforegister.ee where you can search for companies based on their registration code or name.

It is important to show the questionnaire as soon as possible, because the visitor's time on the site might be brief. And the questionnaire should get high attention so that the user feels more inclined to take the survey. This is mostly influenced by where the questionnaire is hosted, in case of platforms that have faster response time are better.

3.3.3 Designing a Delivery Method for the Questionnaire

For the first questionnaire question we had: "What company has a higher reputation?" On visiting a company's page (needs better wording), the application checks for a questionnaire related to that specific company. Once a questionnaire is provided, a modal will be displayed with the questionnaire. The user can then choose one company out of two companies listed. After answering the questionnaire, the user will see statistics about that pair's performance, which out of those two was chosen more often. A prompt is displayed asking the user whether they want another questionnaire. This cycle continues until the question pair for the visited company has been exhausted.

Visual design was heavily inspired by Google's AB testing modal. The modal featured a minimalistic approach to a binary choice questionnaire, with a question and two buttons with a choice's value.

Figure 1. Questionnaire content card

4 Technical Architecture

4.1 Software Selection (Development Tools, Languages)

In my professional career I have developed software using JavaScript and one of the development frameworks called Node.js. Another reason was the compatibility with the current web platform and integration of front-end and back-end services. Therefore, due to my extensive knowledge using this technology, it was chosen as the approach to this thesis's practical side's development.

Node.js has libraries and modules for both front-end and back-end development. Due to this, we can develop our front-end and back-end code using the same programming language. Also, this improves maintainability.

The front-end service is meant as the method of interaction with the visitors. We display the information coming from the back end to the visitors and send the interaction results to the back end for storage.

The back-end service is used to load in data from the database about the specific company's pairs, all the related companies' data and finally the visitor's data in case the visitor is a recurring visitor. Secondly, it serves the front-end views based on the URL the user is requesting. Thirdly, it processes all the data the front-end sends based on the visitor's choices and stores them in the database.

The database is used to store all the data about the companies, all the generated pairs and visitor data. The analysis is performed by data analysis libraries used by codebase created with Python programming language.

4.1.1 JavaScript

JavaScript is a versatile language that can be used for a variety of purposes, including front-end and back-end development. It is one of the most popular programming languages in the world, with a massive community of developers and resources available online. This popularity means that there are many tools, libraries, and frameworks available that can help developers create high-quality applications quickly and efficiently. Also, it is compatible with all modern web browsers, which enables applications to work seamlessly across different platforms and devices. It is also compatible with many popular databases and server-side technologies, enabling it to be used to build full-stack applications.

JavaScript is a high-performance language that can handle enormous amounts of data and complex operations. This makes it well-suited for creating complex web applications that require fast response times. It is a language with a short learning curve, especially for developers who are familiar with other programming languages. It has a simple syntax and is easy to debug.

4.1.2 Node.js

Node.js is a framework that is based on the JavaScript programming language. It was chosen as the main tool to build the solution because of the large library of available modules and how quickly we can build a working concept.

In this case I felt that Node.js is the right tool to use, because we only have a need for a simple front-end UI, back-end, and a basic database. It is also used for creating and setting up the database. Therefore, we have both the front- and back-end of the solution both in similar code that were also executed within the same process, making the code easily understandable and the solution to be as minimal as possible. This enables us to host the code with minimal resources while also enabling high performance.

4.1.3 Front-end Creation

The front-end part of the code is for displaying the questionnaire and the functionality of showing what the back-end sends to be shown. Frontend was written using Pug/Jade and JavaScript, the styling was handled with CSS stylesheet language.

The main requirement is that the questionnaire would be simple, with only 2 buttons for each of the comparable companies and that after the first pair it would have the option to compare another pair.

The front-end must first determine what company the user is visiting and send that company's reg. code. to the back-end, then receive the pair data from the back-end, display it and send the result to the back-end.

Front-end should also use cookies to remove the already compared pair from the options. Another importance of the cookie is to remember if the user has closed the questionnaire by hand, taking it as a sign not to bother the user with the questionnaire for some time.

4.1.4 Pug (Jade) Template Engine

Pug is a template engine for creating clean, organized, and scalable front-end applications. Pug uses indentation to define blocks of code, making it easier to read and maintain.

Pug supports mixins, which allows developers to create reusable code snippets that can be easily integrated into various parts of the application. Another feature includes, with it, developers can split large code files into smaller, more manageable pieces. There are a variety of plugins and extensions available for Pug that can help developers add functionality to their applications.

Pug is highly compatible with other front-end tools and frameworks, such as Angular, React, and Vue.js. This means that developers can easily integrate Pug into their existing workflows and use it alongside other tools to create robust and scalable front-end applications.

4.1.5 Back-end Creation

The back-end application's purpose is to handle the business logic of the questionnaire. The back-end was written using JavaScript programming language to match the front-end. The main framework used is Node.js and the most important packages used are ExpressJS, web framework for Node.js, and Mongoose Object Data Modeling library for MongoDB database and Node.js. Next, we describe in more detail the functionality of the back-end.

Questionnaire related data is fetched and stored in the database. When front-end sends a registration code of a company it fetches a group of pairs for that company. Along with the registration code the front-end also sends the visitor's unique ID from the cookie assigned to the visitor, which relates to that visitor's data in the database, with the information what company pairs the user has answered. The visitor's answered pairs are filtered out from the group of pairs that will be presented and from the resulting group one pair is chosen randomly. Then, using the company registration codes included in the company pair document, company data is fetched from the database related to both companies, such as name. After all the data about each data has been gathered the back-end creates a JSON formatted data object that will be sent to the front-end application.

When a user has chosen a company from the two the front-end sends the pair's ID and which one of the two was chosen. Back-end then sends a message to the database about what the pair ID is and which one of the company's choice counts will be incremented. Also, the visitor's data in the database will be updated with the chosen pair ID so that next time that pair would not be asked from that user.

When the back-end receives a message from the front-end that the visitor closed the questionnaire it then updates the visitor's database record's state to inactive and adds the timestamp of the event. When the same visitor visits the site then the back-end checks the previously recorded data about the visitor, continues with the process like mentioned before when the visitor's state is active, but when it is inactive then checks the timestamp and if two weeks or more has passed then the visitor is set to active again and they will see the questionnaire. Otherwise, the user will not see a questionnaire until their state is set to active.

It is vital that we process the data received from the front-end by data validation and sanitization. By doing so, we ensure that the data is correct and secure. In addition to this, we require that the data is in the correct format and that it is not malicious. Data that does not pass the validation and sanitization process is rejected.

4.1.6 MongoDB Database

MongoDB is a NoSQL database designed for storing and retrieving large amounts of data quickly and efficiently. MongoDB is a document-based database, it stores data in a flexible, JSON-like format that can be easily queried and manipulated. MongoDB is highly scalable, able to handle large amounts of data and high traffic loads without sacrificing performance. MongoDB is also highly available when using cluster solution, which enables it to continue to operate even if some of its nodes fail. It is highly flexible, which means it can be easily integrated with other tools and frameworks to create robust and scalable applications. MongoDB is highly compatible with other tools and frameworks, such as Node.js, Express.js, and Mongoose. It is highly secure, meaning it can protect sensitive data from unauthorized access and ensure that it is stored and transmitted securely. MongoDB is highly reliable; it can ensure that data is stored and retrieved accurately and consistently. MongoDB is highly performant; it can handle large amounts of data and high traffic loads quickly and efficiently.

MongoDB is used as the database for the application created for the questionnaire platform. It consists of simple collections that suit the needs of this research. The benefits we get are the speed for accessing the data and how well it integrates with Node.js. This makes the code minimal and fast to develop, because it does not have the complexity of other database services, for instance PostgreSQL. Another benefit of MongoDB is that it can be set up in a lot of services and at the time of writing there are free options. This makes it great for prototyping and for simpler research cases. There are limitations for free versions, but for the current case they are enough. More about setting up the solution in another section.

The main requirements for the database are:

- Ability of having many collections.

Collections to have many records without any impact to the speed (at least 200 000 records for the pairs, at least 100 000 records for the company data and 500 000 records for visitor data).

- Simple filtering and search.

Records can be simple or complex (records consisting of JSON formatted objects and arrays).

- Importing and exporting of the collections or the whole database.

The database consists of collections and each collection has documents. In this case, the collections are the companies, company pairs and visitor data. The documents here refer to every value in these collections. Scripts for the collections and document creation were created using Node.js. For this we load in the company data with a JavaScript code, thereafter, process the data and create the collections using JSON format.

4.1.7 Python

Python programming language has become one of the most popular programming languages for data analytics and machine learning. Python was chosen for this thesis because of the following reasons: Python has a rich collection of libraries specifically designed for data analysis, such as NumPy, Pandas, and TensorFlow etc. Libraries that provide powerful tools for data manipulation, analysis, and visualization. Python is a flexible and scalable language that can handle large datasets and complex computations. Python has a large and active community of developers and data scientists who contribute to the development of new libraries and tools for data analysis. Python was used to create the codebase for data processing and analysis of the collected questionnaire results.

4.2 Coding Practices and Deployment

In this section, we will discuss the coding practices and deployment methods used in this thesis. We will also discuss the benefits and drawbacks of these methods and the potential impact on the software's quality and performance.

4.2.1 Overview of Coding Methods

4.2.1.1 Clean Code

Clean code is a set of coding practices and principles that help improve the code's readability and maintainability. Some of the key principles of clean code include using meaningful variable names, writing clear and concise comments, and following a consistent coding style. By following these principles, we have created a codebase that is easier to understand and maintain, helping us to reduce the time and effort required to make changes to the code in the future. This also makes the code more reliable and less prone to errors, which can help to improve the overall quality of the software.

4.2.1.2 Refactoring

Refactoring is the process of restructuring existing code to improve its readability, maintainability, and performance. This can involve making changes to the code to remove duplication, improve naming, and simplify complex logic.

4.2.1.3 Newest Technologies and Libraries

Using the newest technologies and libraries can help to improve the quality and performance of the code. By using the latest tools and libraries, we have created a codebase that takes advantage of the latest features and improvements, helping us improve the overall quality of the software.

4.2.1.4 Cloud Services for Deploying and Hosting Codebase

Using cloud services for deploying and hosting the codebase can help to improve the scalability and reliability of the software. By using cloud services, we have taken advantage of the latest infrastructure and tools, which have helped to improve the overall performance and reliability of the software. The main cloud services used were Heroku and mLab. Heroku was used for hosting the codebase, while mLab was used for hosting the database.

4.2.1.5 Server-side rendering

Server-side rendering is the process of generating HTML on the server and sending it to the client. In this thesis, server-side rendering was used to improve the performance and reliability of the software, reducing the amount of work needed to be done on the client side.

4.2.1.6 Git version control

Git version control is a system for tracking changes in codebase files and coordinating work on those files among multiple people. In our codebase it is primarily used for source code management, but it can be used to keep track of changes in any set of files. As a distributed revision control system, it is aimed at speed, data integrity, and support for distributed, non-linear workflows. Git services such as GitHub and Bitbucket were used to host the codebase and track changes in the codebase.

Github was eventually chosen as the main version control system for the codebase due to its popularity and the fact that it is free to use for public and private repositories. It has many features that make it easy to use and it is widely used by developers around the world. Some of its features include issue tracking, pull requests, and code review. GitHub actions were used to automate the deployment process of the codebase. Secrets were used to store sensitive information about the codebase, such as API keys and database credentials.

4.2.2 Serverless and Automatic Deployment

4.2.2.1 Heroku

The questionnaire application developed for the thesis was hosted in web-hosting service Heroku. At the time of hosting the platform provided a free tier. The free tier provided the required amount of storage to host the solution - 50 MiB used out of the available 500 MiB.

The deployment process entailed having the codebase on GitHub and granting the required access to the source control. In addition, Heroku monitored the source code of any updates which activated application deployments. Once the first deployment had been completed a link to the application was provided.

The negative side of Heroku is that after 30 minutes of inactivity the solution is removed from the server's cache. This causes a long initial page load due to cold start and thus visitors might leave the page during that time. Another problem is that only one concurrent user can request from the service and thus people must wait to interact with the solution. Despite these limitations it was still a great hosting solution for our solution.

4.2.2.2 mLab

Our questionnaire application was hosted on Heroku, which provided additional add-ons, including a free hosting add-on for MongoDB called mLab. Adding a mLab database was quick and easy. They provide 512 MiB storage room, which for our study is sufficient - 120 MiB needed for hosting the initial data and by our calculations 100 000 new documents in a collection takes around 25 MiB of storage space. Therefore, we can expect to house at least 1 600 000 documents. In case we exceed the storage, we can export the collected documents and clean the collection of all the documents.

The only shortcoming of the mLab free tier is that it is missing redundancy. In our case weekly backups were sufficient. Also, mLab has a backup scheduling possibility.

mLab has all the MongoDB options for importing and exporting data. Choices are whole database dump, specific documents in a specific collection, whole collection.

The import functionality makes it possible to create the database on our own machine and import it to our mLab when it is ready for the solution. This means we can make quick changes to the database when needed.

The primary collection we need to periodically extract from the database is the company pair documents collection, for this provides the results of the questionnaire.

It is vital to replace the default database accounts with our own accounts. This is done to ensure that the database is secure and that only authorized users have access to the database.

4.3 Codebase management

The main difficulty is devising a questionnaire style and the development of the front-end, back-end and architecting the database. Also, the main challenge was to create a questionnaire that would be easy to use and would not require any complex input from the user. Updating the codebase was a challenge because we had to make sure that the code was up to date and that it was working without any issues. In addition, we had to make sure that the code was secure and that it was not vulnerable to any attacks. Any changes to the codebase had to be tested thoroughly to ensure that they did not introduce any new issues.

Having our own database meant that we had to take care of the data. We had to make sure that the data was continuously backed up and restored when needed. We also had to make sure that the data was not lost and there were measures put into place to recover it. We had to make sure that the data was not corrupted. We had to monitor the database storage to make sure it was not exceeded. These issues could have resulted in a loss of data or would have led to invalid analysis results. We had to make sure that the data was secure and that it was not accessible to unauthorized users. We had to make sure that the data was encrypted and that it was not accessible to hackers.

Hosting our page on the Heroku servers meant that the URL endpoint was unrecognizable for normal users in case we hosted the application on a custom domain. Thus, the questionnaire might have had less visitors or respondents than could have been if the link had been more recognizable. Also, the Heroku free tier has limitations on the number of concurrent users that can access the application. This could have led to a slow response time for the users and could have resulted in poor user experience.

The benefits of having our own codebase are that we can make the changes to the code on the fly, and we can add additional functionalities that we require for our analysis. We can also configure the solution to be modular to embed the questionnaire to another website. This functionality was missing from the questionnaire platforms mentioned in paragraph 3.2.3. In addition, having our own codebase means that we can gather the necessary data for analysis, and we have control over the data. We can access the data when needed and change the values when requirements change. We can expand the database collections when wanted and add new documents or delete documents when needed.

Dangers of having your own database include the risk of invaders or hackers, in case the application or the database are not well protected or continuously upgraded to the latest security patches and standards. In our case, we did not contain any sensitive information, nor was it accessible to users or visitors.

4.4 Data Management

4.4.1 Database Overview and Data Collection

NoSQL database technology MongoDB was used for the thesis to store data provided to the front-end and collected from interaction. We have three collections in the database: companies, company pairs and visitor data. Companies' collection contains the company's registry code, company's name. Company pairs collection contains company 1 registry code, company 2 registry code, company 1 chosen value, company 2 chosen value. Visitor data collection contains the visitor's unique ID, list of company pair IDs.

4.4.2 Data Collected

Data collected for the questionnaire and stored in the database includes:

- Data regarding cookies and users who have answered.
- Company pair ID answered by the user, based on answer the company number 1 or 2 in the company pair chosen value will be incremented.

4.4.3 Data Retrieval Process

The process of retrieving the data for the questionnaire and collecting the result is as follows:

1. Website provides Front-end Application with visited company's registration code,
2. Front-end Application requests a company pair from Back-end Application,

3. Back-end Application fetches all company pairs from Database with matching registration code,
4. Back-end Application chooses randomly one pair from the fetched list and returns it to Front-end application,
5. Front-end application creates a questionnaire that will be displayed to the user,
6. Front-end application records the choice and sends it to Back-end application,
7. Back-end application records the choice to the Database.

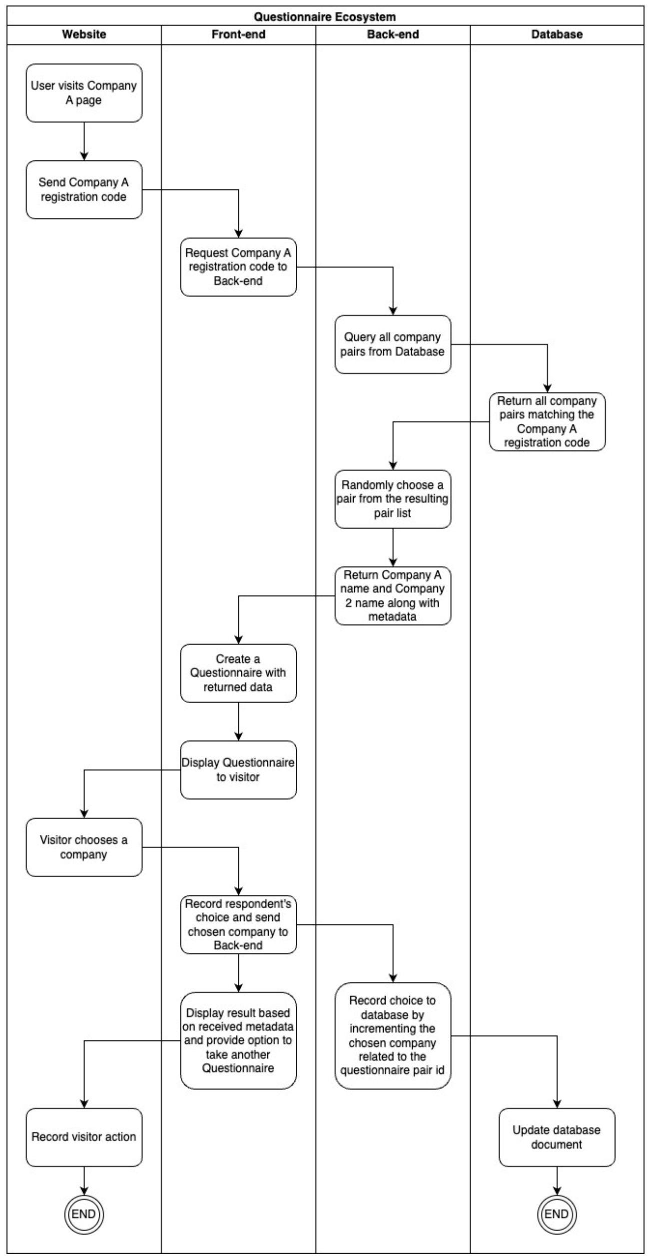

Figure 2. Data retrieval business process diagram

5 Datasets

5.1 Company Dataset

The data collected necessary to create the questionnaire's company pairs are the company’s registry code, name, EMTAK code and financial statistics. The financial statistics include the company’s revenue. The data is collected from the Estonian tax office’s public documents for the year 2018 and includes approximately 120 000 companies and RIK (Registrite ja Infosüsteemide Keskus) including approximately 340 000 companies. The data is collected in a CSV file, which is then processed by the codebase created in Node.js, and finally stored in a MongoDB database. The data is stored in three collections: companies, company pairs and visitor data. The company's collection contains the company’s registry code and name. The company pairs collection contains the company’s registry code, name and the results of the questionnaire. The visitor data collection contains the visitor’s unique ID, the answered company pair IDs. The data is collected for the purpose of creating the questionnaire and analyzing the results to predict the reputation of a company based on the user's choices.

5.2 Field of Activity (EMTAK)

The EMTAK code is a code that indicates the industry field where the company belongs. Using this we can match companies with counterparts belonging to the same field. Through this we increase the validity and the fit of comparison, by comparing companies operating only the same industry field.

5.3 Company list

At the time the thesis was executed there were around 120 000 different companies registered in Estonia. Some of them have stopped existing, and in our case, we are interested in the companies that are still active. In addition, sufficient data about the companies is needed: financial statistics, EMTAK (industry field), name and registry code. The resulting number of companies is 89 000.

In the context of this thesis the method for measuring company reputation makes it possible to find reputation for every company in the context of Estonia. With enough measurements to similar companies, the model becomes more accurate.

5.3.1 Company Data Before Filtering and Combining

The data collected from the Estonian tax office’s public documents for the year 2018 includes information based on which we can filter out the companies that will not add any meaningful value to our thesis and that could invalidate the results. This includes companies that are not active or missing any of the required data. To improve the quality of the data and the results, companies that do not meet the criteria will be filtered out. For this we filter out the companies that are not active, by checking the status of the company in public documents. In addition, companies are removed from the list that are missing any of the required data: financial statistics, EMTAK code, name and registry code of the company. Public data can contain fields that are not relevant to the thesis, such as the company’s address, contact information, etc. These fields are not used in the thesis and are not included in the data.

5.3.2 Combining Company Data

We have details about our company's financial data, EMTAK, company names and their registry code. Next, we will need to combine the data, the key being the registry code. To combine this data, we use Node.js to read in all the data and add the details found to corresponding companies, finally creating a csv file with the resulting complete company data list.

The method we used to combine the company data is as follows: we read in the data from the financial data and the business registry data. Then we filter out the companies that are missing any of the required data. Then we combine the data based on the registry code. We use the registry code as the key to combine the data. We then create a csv file with the resulting complete company data list.

5.4 Generating Comparison Pairs

The method for creating the pairs is as follows: we divide all the companies into groups based on their EMTAK code. Then we divide the companies in the same EMTAK field into three separate groups based on their financial performance (gross amount). Then we create pairs for all these companies, but they will only have a certain set of similar companies as pairs. The pairs are created in a way that the first pair is the company itself and then the next pairs are the companies that are similar to the company. The reasoning behind having different pair creation is to create enough pairs for the comparison and not to create too many. If we had too many pairs, then the results would not create a measurable and repeatable event. Meaning, if the measurement is accurate then we can predict or assume the reputation for other companies thus giving us a good model of how company reputation works.

6 Data analysis and Results

Company reputation is one of the key elements of success and future. There have been many methods for measuring company reputation. During the last decades, the number of computer-based tools and solutions to measure company reputation has rapidly increased. However, easy tasks become more complex when different components of qualitative and quantitative research data are considered.

As Lange et al. argue the reputation is rooted in the organization's historical behavior and associations but can be abruptly changed if new information about the organization's past behavior becomes known or if the organization's latest behaviors or associations are jarring to observers. They emphasize the diversity of definitions for the organizational reputation construct that are evident in the management literature, and describe conceptualizations of organizational reputation: being known, being known for something, and generalized favorability (Lange et al. 2011, pp. 155).

6.1 Data Collection and Management

6.1.1 Distributing the Questionnaire

Primary focus when creating the questionnaire was making an accessible and coherent user interface with clear functionality and purpose. It is imperative that the visitor knows directly what the question is and what are the possible options to choose. For this, the UI was designed as a content card with a question text, followed by two images and descriptions of the companies that when pressed would correspond to the name of the company. There would also be a possibility to close the card, which would then dismiss the questionnaire and pass along a request to the application that the visitor would not see the card for a specified period. The other important functionality was to show the results of the questionnaire pair in question. The goal behind the result is to show more interesting facts instead of solely a thank you message. This makes the questionnaire more informative and adds incentive to the visitor to answer more questionnaires.

As mentioned previously, the questionnaires display two companies, one of which is based on the company visitor is currently displaying, and the other will be fetched from the database based on the currently relevant company. The visitor will not see a questionnaire if there are no possible questionnaires to be shown. This includes the case when the visitor has already answered all the possible questionnaires for the company. This functionality was lacking in other questionnaire platforms. This is due to lack of input dynamic questions, which in our case is a major drawback as it limits the possibilities of creating a more engaging and informative questionnaire.

For the experiment we needed a set of companies to be displayed to potential visitors or visitors who helped with the thesis. We chose 23 companies from Äripäev Revenue TOP 500 companies list based on financial performance and the number of company pairs we had created with companies running in the same field of industry. Companies chosen were following: Olympic Entertainment Group AS, ABB AS, Stora Enso Eesti AS, Silberauto AS, Eesti Energia AS, LHV Group AS, Orkla Eesti AS / Kalev AS, Sportland International Group AS, Riigimetsa Majandamise Keskus, Tallinna Vesi AS, Tallinna Kaubamaja Grupp AS, Eesti Loto AS, Nortal AS, ABC Grupi AS, Liviko AS, SEB Pank AS, SWEDBANK AS, Lennuliiklusteeninduse

Aktsiaselts, Ekspress Grupp AS, Rimi Eesti Food AS, Tallinna Linnatranspordi AS, RAGN-SELLS AS, DHL Estonia AS.

Initially the questionnaire platform was developed for the company data portal Inforegister.ee, which enabled integration with direct company data necessary to provide questionnaire matching the precise context. A context aware questionnaire design was the key to our research, removing the randomness of questionnaire and providing accurate data for further analysis.

For the initial launch of the questionnaire, an actual company data platform should not be used due to the following reasons:

- Visitors may take questionnaires regarding companies that are not measured in typical reputation lists resulting in irrelevant results, which can create too much noise in the data.
- Too large data flow cannot be handled by the initially developed solution, a more costly and high availability application would need to be set up, however this would not improve the results.
- Any updates to the application would need to be coordinated with the websites where the application has been integrated to.
- Limitation of Heroku, mLab – limited database size, low server availability, unforeseen errors.

To collect statistically relevant results for our questionnaire, an environment was developed to integrate the questionnaire application to simulate ecosystem that the application would be integrated into. This way we can provide tasks for the visitors that would provide data for our analysis that can be compared to results of other company reputation lists (Äripäev TOP 100).

To distribute the questionnaire, we created a static website that listed the top 23 companies. The landing page had a brief introduction to the questionnaire followed by the list of the aforementioned companies. The visitor was instructed to select a company from the list and answer the questionnaires created for that company. The visitor was encouraged to answer as many questionnaires as they could. After exhausting all the possible company pairs, a dialogue was presented to indicate that there were no more questionnaires left, after which they were redirected back to the front page helping the process of choosing another company.

The static website was hosted in Heroku along with the questionnaire application. The link to the questionnaire was distributed through friends, family and Facebook.

6.1.2 Collecting the Results

To measure the choice between two companies we recorded the answer from the visitor in our database. The method was to collect the identification of the company pair that was displayed to the visitor and which one of the two companies was chosen by the visitor. Then the result was passed to the back-end which modified the database document by incrementing the matching questionnaire pair's chosen company's value. Eventually, we had a measurable result indicating the companies chosen as the most reputable companies to the visitors.

Along with storing the company which the visitor found more reputable out of a company pair list, the application also stores the list of company pairs the visitor has answered. However, this list is not used in the analysis, but it is used to prevent the visitor from seeing the same company pair again. This is to prevent the visitor from answering the same question multiple times, which would skew the results.

6.1.3 Collected Data

The questionnaire was presented on 30th January 2020 and was available for 2 weeks. The resulting collection of questionnaire results was 1160 answers. For all the answers in the pair data collection, each answer was treated as a new record, indicating whether the first company in the company pair was chosen, represented in binary format: 0 or 1. To extract the collected data a Node.js codebase was developed. The application created a connection to the database, it then queried all the documents from company pairs collection and finally wrote each document into a CSV file. This step was followed by data preprocessing.

The following is an excerpt from collected data with all the included data columns:

Table 1. Excerpt from Collected Data

Company 1 Chosen	Company 2 Chose	Company 1 Name	Company 2 Name	Company 1 Turnover	Company 2 Turnover	Company 1 Registration Code	Company 2 Registration Code
1	0	LHV PANK AS	LUMINOR BANK AS	33638000	12397	10539549	11315936
0	1	LHV PANK AS	LUMINOR BANK AS	33638000	12397	10539549	11315936
0	1	LHV PANK AS	LUMINOR BANK AS	33638000	12397	10539549	11315936
1	0	LHV PANK AS	INBANK AS	33638000	6039000	10539549	12001988

6.1.4 Data Preprocessing

In this section, we describe the data preparation for training the machine learning models. Data preprocessing is an essential step in preparing the dataset for training machine learning models. The goal of data preprocessing is to clean, transform, and scale the data to ensure that it is suitable for training and evaluation. The steps involved in data preprocessing include data cleaning, checking for missing values, encoding categorical variables, and feature scaling.

We need to process the data from the questionnaire into a dataset to be trained on. The dataset was created from the results of the questionnaire. The columns of the dataset are the company chosen from the company pair, first company's registration code, second company's registration code, first company's financial performance and the second company’s financial performance.

After the questionnaire results were gathered, the data was extracted from the remote database into a local copy. In the company pair table, we recorded which company was chosen incrementing the value of the chosen company - either company one or company two. The data was processed by recording each questionnaire choice from the questionnaire pair as its own row. Each choice has a Boolean indicator, representing whether the first company from the questionnaire pair was chosen. In the questionnaire pair table, we did not have any information about the company's financial performance, for this we had to match the company's registration code with the value from the company data table. As a result, we had a table containing each of the answers given by the visitors, with an indicator for which company from the company pair was chosen. The table was written into a csv file with each row's data processed based on its type and then added into a column. This format is needed for creating the model.

In data cleaning step we will remove any irrelevant or duplicate data from the dataset. This will help to ensure that the dataset is clean and free of any errors that could affect the performance of the models.

Checking for missing values is to ensure that there are no missing values in the dataset. Missing values can introduce bias in the model and affect the accuracy of the predictions.

In case the company names are categorical, we will need to encode them into numerical values using techniques like one-hot encoding. This will help to ensure that the models can process the data and make accurate predictions. For instance, we can encode the company names into numerical values using one-hot encoding.

In addition to the aforementioned steps, we will also need to scale the features to ensure that they are on a similar scale. If the numerical features (e.g., financial performance) are on different scales, the values need to be scaled to a similar range to avoid bias in the model. Models like logistic regression and support vector machines are sensitive to feature scaling while decision trees and random forests are not. In this case, we will scale the features to a similar range to ensure consistent performance across different models.

Feature scaling was done to ensure that the numerical features (e.g., financial performance) are on the same scale. This is important to avoid bias in the model and ensure that all features contribute equally to the prediction.

Database extraction and data processing was executed by codebase created using JavaScript programming language and Node.js framework, because we used a similar codebase for creating the company list and company pair list. For analysis we created a codebase in Python using the following libraries for executing the analysis: Pandas, SKLearn and Numpy. The table containing the results was read by Pandas from a CSV file. Then a dataset was created from the data rows, each row's data processed based on its type and then added into a column.

6.2 Experimental Research

In this section, we will discuss the methodology for choosing the model to predict the reputation of a company based on the user's choices. We will explore the different classification algorithms that can be used to analyze the collected data and predict the reputation of a company. We will also discuss the interpretability of the models and how they can provide insights into the factors that influence the choice of a company. Finally, we will outline the steps for preparing the data, training the models, and evaluating their performance.

6.2.1 Feature Selection and Engineering

Feature selection and engineering are important steps in improving the predictive performance of machine learning models. The goal of feature selection is to identify the most relevant features for prediction, while feature engineering involves creating new features from existing ones to improve the model's performance. Using techniques like correlation analysis or feature importance from tree-based models we can identify the most relevant features for prediction. Creating new features from existing ones could provide additional predictive performance.

6.2.2 Model Selection

Model selection is a critical step in choosing the best algorithm for predicting the reputation of a company based on user choices. The choice of model depends on the dataset characteristics, the complexity of the relationships between features, and the interpretability of the model. The steps involved in model selection include choosing a set of candidate classification algorithms and training multiple models. In the present study we will be training and evaluating the resulting models for following algorithms: logistic regression, decision trees, random forests, support vector machines and neural networks.

The collected data from the questionnaire can be analyzed using various machine learning models to predict the reputation of a company based on the user's choices. The models can be trained on the collected data to learn the relationship between the features and the outcome, helping us predict the reputation of a company based on the user's choices. The models can be evaluated based on their performance metrics, such as accuracy, precision, recall, and F1 score, to determine the best model for predicting the reputation of a company. In this section we will discuss some of the algorithms and machine learning models that can be used to analyze the collected data.

Interpretability is an important aspect of machine learning models, especially in applications where understanding the decision-making process is crucial. In the context of predicting the reputation of a company based on user choices, interpretability can provide insights into the factors that influence the choice of a company. Different classification algorithms have varying levels of interpretability, which can impact their utility in understanding the relationship between features and the predicted outcome.

Logistic regression is a simple and widely used linear classification algorithm. It models the probability of a binary outcome based on one or more predictor variables. Logistic regression can be suitable for our dataset if the relationship between the features and the probability of choosing the first company is approximately linear or can be adequately captured by a linear decision boundary. It is interpretable and computationally efficient, making it a desirable choice for initial exploration or as a baseline model. Logistic regression can provide insights into the relationship between the features and the outcome, helping us understand the factors that influence the choice of the first company. Linear decision boundaries may not be able to capture complex relationships or interactions between features, which could limit the model's performance in datasets with non-linear relationships or high-dimensional feature spaces. Logistic regression may not be suitable for datasets with non-linear relationships or complex interactions between features, as it assumes a linear relationship between the features and the outcome. In such cases, more complex models like decision trees, random forests, support vector machines, or neural networks may be more appropriate. Linear decision boundary is a straight-line separating data into two classes.

Logistic regression models are highly interpretable. The coefficients of logistic regression directly represent the relationship between the input features and the log-odds of the target variable. Positive coefficients indicate a positive association with the target variable, while negative coefficients indicate a negative association. The number of coefficients can indicate the importance of features in predicting the target variable. Features with larger coefficients have a greater impact on the predicted probability.

Decision trees are a non-linear, tree-structured model that partitions the feature space into regions, assigning a class label to each region. It can capture non-linear relationships and interactions between features, which may be present in our dataset. They are interpretable and can provide insights into the decision-making process. Decision trees could be useful for identifying important company attributes that influence the choice of the first company. However, they are prone to overfitting, by capturing noise in the data and performing poorly on unseen data. To mitigate overfitting, we can use techniques like pruning or setting a maximum depth. Decision trees may not be suitable for datasets in case of high-dimensional features or complex relationships between features, as they may struggle to capture complex decision boundaries. Decision trees are inherently interpretable. They partition the feature space into a series of simple decision rules based on feature values. Each decision node represents a decision based on a feature, and each leaf node represents the predicted class. Decision trees rank features based on their ability to split the data effectively. Features appearing closer to the root node are more important in making predictions.

Random forests is an ensemble learning technique that combines multiple decision trees to improve performance and reduce overfitting. It can handle complex relationships and interactions between features, making them suitable for datasets with high-dimensional features or non-linear relationships. They are resilient to overfitting and can provide better generalization performance compared to individual decision trees. Random forests can capture complex decision boundaries and are effective for identifying important company attributes that influence the choice of the first company. Random forests may not be suitable for datasets with a small number of features or simple relationships between features, as they may introduce unnecessary complexity and reduce model interpretability. While individual decision trees are interpretable, the interpretation of a random forest is more challenging due to the aggregation of multiple trees.

SVMs (Support Vector Machines) are powerful supervised learning models that find the optimal hyperplane to separate classes in a high-dimensional space. They are effective for binary classification tasks and are particularly useful when the decision boundary is non-linear. SVMs are resilient to overfitting and can provide good generalization performance on unseen data. SVMs may not be suitable for datasets with many features or complex relationships between features, as they may introduce unnecessary complexity and reduce model interpretability, making it challenging to understand the decision-making process. The decision boundary of an SVM is determined by support vectors, which are a subset of training samples. Understanding the decision boundary in high-dimensional spaces can be complex.

Neural networks are a class of deep learning models inspired by the structure and function of the human brain. They consist of interconnected layers of neurons that learn complex patterns from data. Furthermore, Neural networks excel at learning complex and non-linear relationships in data. If our dataset contains high-dimensional features or complex interactions between features, neural networks could be effective. Neural networks may not be suitable for datasets with a small number of features or simple relationships between features, as they may introduce unnecessary complexity and reduce model interpretability. In addition, they are less interpretable, especially as they become more complex with multiple layers and neurons. Understanding how each neuron contributes to the final prediction can be difficult.

6.2.3 Model Evaluation

Model evaluation is an essential step in assessing the performance of machine learning models. The goal of model evaluation is to determine the accuracy and generalization performance of the trained models. The steps involved in model evaluation include dataset splitting into training and testing sets, training models, and evaluating their performance using appropriate evaluation metrics. A portion of 20% of the dataset will be separated for testing to evaluate the performance of trained models.

In this section, we will evaluate the performance of the models using various metrics like accuracy, precision, recall, F1-score, and ROC-AUC score. These metrics will help us determine the best model for predicting the reputation of a company based on the user's choices. We will compare the performance of logistic regression, decision trees, random forests, support vector machines, and neural networks to identify the most accurate and reliable model.

In more detail accuracy is the ratio of correctly predicted instances to the total number of instances in the dataset. Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. Recall is the ratio of correctly predicted positive observations to all observations in actual class. F1-score is the weighted average of Precision and Recall. ROC-AUC score is the area under the receiver operating characteristic curve, which is a plot of the true positive rate against the false positive rate.

Table 2. Evaluation of models

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC score
Logistic Regression	92.2%	93%	99.1%	95.9%	52.5%
Decision Trees	88.8%	92.8%	95.3%	94%	50.6%
Random Forests	90.5%	92.9%	97.2%	95%	51.5%
SVM	92.8%	92.7%	100%	96.2%	50%
Neural Network	90.5%	92.9%	97.2%	95%	51.5%

- Based on analysis results how well the model used fits for the data gathered.
- Which is the most suitable model for our research.
- How well does the best model predict results.

Next, we will interpret the results of each model.

Accuracy measured for each model indicates a good overall performance in classification tasks. Here Logic Regression and SVM have taken the lead, while other models perform slightly less accuracy with room for more improvement.

Precision measured indicates a low chance of false positives for all models.

A high recall results indicate the rate of success at identifying positive cases (true positives). SVM outperforms other models by perfectly identifying all actual positives, which for other models is slightly lower due to dataset imbalances.

F1-score balances precision and recall into one metric and a high value suggests good performance for both aspects. All the models show an overall good performance, with SVM showing a higher value than others, which might require further investigation into dataset characteristics or potential overfitting.

A common low resulting ROC-AUC Score indicates difficulties in distinguishing classes effectively or an imbalanced dataset.

To further evaluate the model, we want to know if the result was in any way affected by the Halo effect (Brown and Perry 1994, pp. 1349), whether the result was dependent on the financial performance of the company chosen. In this case we assume that the company chosen has a higher financial performance compared to the other company. For this, we gathered the results of all the questionnaire pairs, then checked for the scale of companies chosen and had the higher financial performance of the two. As a result of the question, we also got 92.24%. This means that the model trained has learned to base the result of the company pair comparison to the financial performance of the companies, thus making the same mistake when a company with a smaller financial performance was chosen.

6.3 Analysis of Results

6.3.1 Prediction

Through our experimental research, we have determined the performance of multiple models to predict other company reputations. In the case that our data is imbalanced, Decision Trees or Random Forests can be the best choices. However, due to the large volume of companies and additional features being added in the future, Neural Networks will prove to provide better results. In addition, Neural Networks are better in cases where positive cases are critical.

However, SVM has provided better overall results in the model evaluation making it the best model for future predictions for company reputation in case of additional companies. This is due to SVM resulting with the highest accuracy and F1-score. SVM balances precision and recall effectively.

6.3.2 Validity

In this section we describe the model's validity, determined by the accuracy of the performance metrics described in model evaluation section 6.2.3. The model is considered valid if:

- the predictions are accurate and reliable,
- it provides insights into the factors that influence the choice of a company,
- the performance metrics are high and consistent across different evaluation metrics,
- it can be used to predict the reputation of a company based on the user's choices,
- the predictions are close to the actual values.

In the case of the models evaluated we can consider all of them accurate and reliable due to the Accuracy and Precision results. The models were given financial performance as one of the features and the results give us a clear insight into the effect financial performance has compared to the choices questionnaire respondents provided. In our case we were not able to publicly gather any additional features, therefore we can assess the results to reflect the financial performance as the main attribute the respondents used to make their choices.

7 Discussion

7.1 Interpretation of Findings

A rigorous analysis of related works was conducted to find the criteria used to measure the company's reputation. For our research the most descriptive research question was found to be “What company has a higher reputation?”. This was also considered by Ponzi, Fumbrun and Gardberg in their 2011 RepTrak® Pulse instrument, along with Lange, Lee and Dai in their 2011 “study on mobile phone providers”.

In our study’s experimental research, we first established the source of publicly available company data which is the cornerstone for building the list of companies that we will measure the reputation for. We developed applications to gather the company data from multiple sources, which then processed the data into a list of companies grouped by their industry fields to provide company pairs to a custom developed context aware questionnaire platform. This platform was integrated into a questionnaire website to collect results from the respondents. The collected data enabled further model development and analysis. We trained multiple models comprehensively analyzing each model's results by their accuracy, precision and additional metrics.

In our research we conclude that the ecosystem developed for measuring company reputation is capable to create many questionnaires that were presented to multiple respondents, resulting in data that provides significant results to measure company reputation. This result was validated by the models trained on the data.

7.2 Comparisons to Existing Literature

In their research, Fombrun et al. (2000) found eight attributes; in this study, one attribute was used for the research and analysis. However, the application can provide additional attributes derived from tools such as RepTrak®, AMAC, and RQ to research the halo effect, objective and subjective results, tangible and intangible assets, and qualitative and quantitative data.

Secondly, Fombrun et al. proposed that 'the reputation quotient' (RQ) recognizes the difficulty in developing a valid database of corporate reputational ratings. An extensive amount of development went to the design and architecture of the database to provide a robust questionnaire platform and analysis possibilities. They concluded that Fortune AMAC, released in 1984, can be considered the backbone for most empirical academic work on reputation. Indeed, the criteria of financial soundness and stability have been a centric criterion for Management Today (MT), Manager Magazine (MM), Asian Business (AB), Far Eastern Economic Review (FEER), Financial Times (FT), Industry Week (IW), Fortune GMAC. In addition, financial performance has been the most important to Estonia's Äripäev Revenue TOP 500 and TOP 100 research and financial data of Estonian companies is publicly available. This thesis used financial performance to create the questionnaire comparison pairs.

Thirdly, GMAC (Global Most Admired Companies) interviewed about 8000 people since 1983; in 1985, the survey's return rate was 50%. The potential of our developed questionnaire platform depends on the website's number of visitors. In the case of Inforegister.ee, the yearly amount is around 300 000 people. Integrating our platform into additional websites (Inforegister.ee, e-krediidiinfo.ee and others) provides a marginally higher number of possible questionnaire respondents and resulting responses than other methodologies.

Lange, Lee, and Dai (2011) conducted their research "study on mobile phone providers" by contacting 2200 randomly selected students from all fields at major German universities, out of which 306 respondents completed the survey entirely. This approach of limiting the survey to a specific industry matches our system of separating the companies by their industry field to improve the questionnaire's context. However, our solution enables us to conduct similar studies across all industries simultaneously.

Ponzi, Fombrun, and Gardberg (2011) developed an improved RepTrak® Pulse to rate a company's overall reputation by measuring with a 7-point scale. This scale, however, does not consider the company's industry, which our research considers when comparing companies. The thesis method provides a more comprehensive and industry-focused result, enabling ranking by industry field.

When Caruana (1997) researched the effects of positive and negative halo effects on a company's long term, a positive halo ensures a more stable reputation while a negative halo keeps the company from restoring its reputation even through attempts to restore it. The results from thesis research enables us to chart a historical change in a company's reputation compared to others in the same industry field, enabling us to detect positive or negative halo and its effects on the company's financial performance.

Finally, Schultz et al. (2001) highlight unknowable chances for hunches and guesses from respondents that can depend on the respondents themselves, the question provided, and the methods used. The research proves the subjective nature of responses to the questionnaire, which, in our case, explains the accuracy and prediction results of our trained models.

8 Conclusion and future work

This research presented a comprehensive approach to automating the estimation of company reputation. Our thesis developed upon various methodologies, including questionnaire design, data management, and the use of machine learning models such as Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines, to analyze the perception of company reputation. While we made progress in reputation analysis, we also found potential drawbacks in developing our own methodologies such as maintenance requirements, browser compatibility issues, website integration challenges and data protection policies. The scalability and feature implementation of the questionnaire platform required a constant balance between flexibility and complexity.

In our study of related works, we highlighted the significant influence of corporate reputation on a company's value, affecting investor interest, customer attraction, and retention. The underlying difficulty in measuring corporate reputation relies heavily on the need for objective criteria over subjective assessments.

Our research focuses on the importance of data preprocessing and model evaluation to ensure accurate reputation prediction.

The main contributions of this thesis were:

- Innovative Methodology: Development of a new methodology to evaluate company reputation by combining network analysis and machine learning methods. This allows for the assessment of reputation for companies of any size in Estonia.
- Questionnaire Design: Creation of a unique questionnaire platform, which collects comparative data between two similar companies to determine their reputation. This design includes unique features to enhance user experience and retention.
- Technical Architecture: The thesis outlines a robust technical architecture for the implementation of the reputation estimation system. It includes the selection of software, development tools, languages, and coding practices, with a focus on serverless and automatic deployment.
- Empirical Research: Conducting experimental research and analyses to validate the proposed models and methodologies. Various data analysis models are employed, such as Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), and Neural Networks, to analyze the collected data and predict company reputation.

These contributions are significant as they provide a scalable and objective way to measure company reputation, which is an asset for organizations in understanding their standing with stakeholders and in the marketplace.

8.1 Future research

The thesis opens possibilities for future work, such as refining the machine learning models for greater accuracy and expanding the research to a broader range of companies and industries. Further development of the questionnaire platform to enhance user experience and data collection is also suggested. Further development to refine the reputation assessment model, in addition to exploring its applicability in different contexts, and integration additional variables should be considered for a more comprehensive evaluation. This research focused on the overall reputation of companies, however additional dimensions such as emotional appeal, products and services, vision and leadership, workplace environment, social responsibility (Roper and Fill 2012, pp. 91; Fombrun et al. 2000, pp. 246) can be considered as additional research questions for the questionnaire.

Future research can include a wider range of datasets and consider reputation in the digital era. A valuable source for this data can be found from company's social media, articles and the websites that they maintain.

The models used for reputation analysis can be further developed, particularly in adapting to different cultural contexts and industry field specific nuances. A well-designed ecosystem of data collection, processing and analysis architecture has the potential for real-time reputation tracking.

References

Barchiesi, M. A., La Bella, A. (2014). An Analysis of the Organizational Core Values of the World's Most Admired Companies. In Knowledge and Process Management, 21 (3), pp. 159–166.

Bermiss, Y. S., Zajac, E. J., King, B. G. (2014). How commensuration and management fashion affect corporate reputation rankings. In Organization Science, 25 (2), pp. 591–608.

Brown, B., Perry, S. (1994). Removing the Financial Performance Halo from Fortune's "Most Admired" Companies. In Academy of Management Journal, 37 (5), pp. 1347–1359.

Bruner, J. S. (1957). On perceptual readiness. In Psychological Review, 64, pp. 123–152.

Calderón, R., Ferrero, I., Redin, D. M. (2012). Ethical codes and corporate responsibility of the most admired companies of the world: Toward a third generation ethics? Business and Politics, 14 (4), pp. 1–24.

Campbell, K., Sherman, W. R. (2010). Lists & More Lists: Making Sense of Corporate Reputations. Journal of Business & Economics Research, 8 (7), pp. 47–58.

Cannon, H. M., Schwaiger, M. (2005). The role of company reputation in business simulations. In Simulation & Gaming, 36 (2), pp. 188–202.

Caruana, A. (1997). Corporate reputation: concept and measurement. In Journal of Product & Brand Management, 6 (2), pp. 109–118.

Cherchiello, P. (2011). Statistical Models to Measure Corporate Reputation. In Journal of Applied Quantitative Methods, 6 (4), December 30, pp. 58–71.

Coyne, E. E. (2010). Reputation as Information: A Multilevel Approach to Reputation in Organizations. Dissertation. The Ohio State University.

Cravens, K., Goad O. E., Ramamoorti, S. (2003). The Reputation Index: Measuring and Managing Corporate Reputation. In European Management Journal, 21 (2), pp. 201–212.

Davies, G., Chun, R., Da Silva, R. V., Roper, S. (2003). Corporate Reputation and Competitiveness. London & New York: Routhledge.

Fombrun, C., Shanley, M. (1990), What's in a name? Reputation building and corporate strategy'. In Academy of Management Journal, 33, pp. 233–258.

Fombrun, C. J. (1996). Reputation: Realizing the Value from the Corporate Image. Boston: Harvard Business School Press.

Fombrun, C. J., Van Riel, C. (1997). The Reputational Landscape, Corporate Reputation Review, 1 (1), pp. 5–13.

Fombrun, C. J., Gardberg, N. A., Sever, J. M. (2000). Reputation Quotient model. Reputation management. EURIB, European Institute for Brand Management. Available at: http://eurib.prelive.nl/fileadmin/user_upload/Documenten/PDF/Reputatiemanagement/t_-_Reputation_Quotient_model.pdf

Fombrun, C.J., Gardberg, N.A., Sever, J.M. (2000). The reputation quotient: a multi-stakeholder measure of corporate reputation. In Journal of Brand Management, 7 (4), pp. 241–255. Available at: https://link.springer.com/article/10.1057/bm.2000.10

Fombrun, C. J., Van Riel, C. (2003). Fame & Fortune: How Successful Companies Build Winning Reputations. London: Pearson.

Fombrun, C. J., Gardberg, N. A., Sever, J. M. (2009). The Reputation QuotientSM: A multi-stakeholder measure of corporate reputation. In Journal of Brand Management, 7 (4), pp. 241–255.

Fryxell, G. E., Wang, J. (1994). The Fortune Corporate Reputation Index: Reputation for What? In Journal of Management, 20, pp. 1–14.

Gotsi, M., Wilson, A. M. (2001). Corporate reputation: seeking a definition. In Corporate Communications: An International Journal, 6 (1), 24–30. Available at: http://dx.doi.org/10.1108/13563280110381189 (21.10.2013)

Harvey, M.G., Lusch, R.F. (1999). Balancing the intellectual capital books: intangible liabilities. In European Management Journal, 17 (1), pp. 85–92.

Herbig, P., Milewicz, J. (1995). To be or not to be…credible that is: a model of reputation and credibility among competing firms. In Marketing Intelligence & Planning, 13 (6), pp. 24–33 (10).

Heugens, P. (2004). Reviews the book "Fame & Fortune: How Successful Companies Build Winning Reputations," by Charles J. Fombrun and Cees B. M. van Riel. In Corporate Reputation Review, 6 (4), pp. 390–395.

Jacoby, J., Speller, D. E., Kohn, C. A. (1974). Brand choice behaviour as a function of information load. In Journal of Market Research, 11, February, pp. 63–69.

Jiang, E. (2020). Here are the 21 most brilliant quotes from Warren Buffett, the world's most famous and successful investor. In Business Insider, Markets Insider, 12.02.2020. Available at: https://markets.businessinsider.com/news/stocks/warren-buffett-21-best-quotes-2019-2-1027944381 (23.11.2020).

Jin, B., Park, J. Y. Kim, J. (2008). Cross-cultural examination of the relationships among firm reputation, e-satisfaction, e-trust, and e-loyalty. In International Marketing Review, 25 (3), pp. 324–337. Available at: http://dx.doi.org/10.1108/02651330810877243 (21.10.2013)

Juliusson, G., Eriksson, P. (2011). On the relationship between corporate reputation and brand trust – an investigation of three corporate brands. Gothenburg: University of Gothenburg School of Business, Economics and Law.

Lange, D., Lee, P. M.; Dai, Y. (2011). Organizational Reputation: A Review. In Journal of Management, 37 (1), pp. 153–184.

McGuire, J. B., Schneeweis, T., Branch, B. (1990) 'Perceptions of firm quality: A cause or result of firm performance'. In Journal of Management, 16, pp. 167–180.

Park, N., Lee, K. M. (2007). Effects of online news forum on corporate reputation. In Public Relations Review, 33, pp. 346–348.

Ponzi, L. J., Fombrun, C. J., Gardberg, N. A. (2011). RepTrakTMPulse: Conceptualizing and Validating a Short-Form Measure of Corporate Reputation. In Corporate Reputation Review, 14 (1), pp. 15–35. Available at: https://doi.org/10.1057/crr.2011.5

Roper, S., Fill, C. (2012). Corporate Reputation: Brand and Communication. Pearson Education Limited. Pearson Education; Pap/Psc edition (May 30, 2012)

Puncheva, P. (2008). The Role of Corporate Reputation in the Stakeholder Decision-Making Process. In Business & Society, 47 (3), September, pp. 272–290.

Sabater, J., Sierra, C. (2002). Social ReGreT, a reputation model based on social relations. In ACM SIGecom Exchanges, 3 (1), pp. 44–56.

Sarstedt, M., Schloderer, M. P. (2010). Developing a measurement approach for reputation of non-profit organizations. In International Journal of Nonprofit and Voluntary Sector Marketing, 15 (3), pp. 276–299.

Sarstedt, M., Wilczynski, P., Melewar, T. C. (2013). Measuring reputation in global markets-A comparison of reputation measures' convergent and criterion validities. In Journal of World Business, 48 (3), pp. 329–339.

Sarstedt, M., Wilczynski, P., Melewar, T. C. (2012). Measuring reputation in global markets – A comparison of reputation measures’ convergent and criterion validities. In Journal of World Business, 48 (3), pp. 329–339.

Schwaiger, M. (2004). Components and parameters of corporate reputation – an empirical study. In Schmalenbach Business Review, 56, pp. 46–71.

Schultz, M., Mouritsen, J., Gabrielsen, G. (2001). Sticky Reputation: Analyzing a Ranking System. In Corporate Reputation Review, 4 (1), pp. 24–41.

Serbanica, C. (2009). The perception over corporate reputation, a picture of the country profile. A Romanian 1. A Theoretical Approach to Corporate Reputation Evaluation. In Verslo ir teisės aktualijos/Current Issues of Business and Law, 3, pp. 47–65.

Serbanica, C., Popescu, T. (2009). The Perception Over Corporate Reputation, a Picture of the Country Profile. A Romanian Case Study. In Current Issues of Business and Law, 3, pp. 47–65.

Smaiziene, I., Jucevicius, R. (2009). Corporate Reputation: Multidisciplinary Richness and Search for a Relevant Definition. In Inzinerine Ekonomika-Engineering Economics, 2, pp. 91–101.

Tischer, S., Hildebrandt, L. (2011). Linking corporate reputation and shareholder value using the publication of reputation rankings. In SFB 649 Discussion Paper 2011-065, available at: http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2011-065.pdf (27.10.2013)

Walsh, G., Beatty, S. E., Shiu, E. M. K. (2009). The customer-based corporate reputation scale: Replication and short form. In Journal of Business Research, 62, pp. 924–930.

Walsh, G., Beatty, S.E. (2007). Measuring customer-based corporate reputation: scale development, validation, and application. In Journal of the Academy of Marketing Science, 35 (1), pp. 127–143.

Walsh, G., Wiedmann, K.-P. (2004). A Conceptualization of Corporate Reputation in Germany: An Evaluation and Extension of the RQ. In Corporate Reputation Review, 6 (4), pp. 304–313.

Watson, T., Kitchen, P. J. (2010). Reputation Management: Corporate Image and Communication. Available at: http://eprints.bournemouth.ac.uk/12633/2/Chapter_13_def.pdf (22.10.2013)

Williams, R. J., Barrett, J. D. (2000). Corporate Philanthropy, Criminal Activity, and Firm Reputation: Is There a Link? In Journal of Business Ethics, 26 (4), pp. 341–350.

RIK (2019). Ettevõtja Rekvisiidid 2019. Available at https://avaandmed.rik.ee/andmed/ARIREGISTER/ accessed 14.08.2019

EMTA (2019). Company Taxes paid, turnover and size of workforce in Q1 of 2019 https://www.emta.ee/en/private-client/board-news-and-contacts/news-press-information-statistics/statistics-and-open-data accessed 14.08.2019.

Kantar Emor (2008). Eesti kõige mainekam suurettevõte on Hansapank. Available at https://www.kantaremor.ee/pressiteated/eesti-koige-mainekam-suurettevote-on-hansapank/ accessed 10.08.2019.

Kantar Emor (2008). Emori uuring: Eesti mainekaim suurettevõte on Rakvere Lihakombinaat, eelistatuim tööandja aga Eesti Energia. Available at https://www.kantaremor.ee/pressiteated/emori-uuring-eesti-mainekaim-suurettevote-on-rakvere-lihakombinaat-eelistatuim-tooandja-aga-eesti-energia/ accessed 10.08.2019

Äripäev (2017). Äripäev Revenue TOP 500 visited 22.09.2019.

Appendix 1 – Non-exclusive licence for reproduction and publication of a graduation thesis[1]

I Koit Summatavet

1. Grant Tallin University of Technology free licence (non-exclusive license) for my thesis "Automated Estimation of Company Reputation", supervised by Innar Liiv
   1.1. to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   1.2. to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

Tallinn 12.05.2024

1 The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.