

University of Camerino

SCHOOL OF SCIENCE AND TECHNOLOGY
Master Degree in Computer Science (LM-18)



**Design and Implementation of Reputation
System in OSINT Environment**

Candidate
Gian Marco Mennecozi

Advisor
Prof. Fausto Marcantoni

Co-Advisor
Ing. Gaetano Ascenzi

A.Y. 2018/2019

Abstract

Reputation is one of the many intangible assets of a company, an asset that cannot be bought. Having an excellent reputation means being considered reliable and therefore portraying a positive social image. A good reputation is an intangible strategic asset, capable of generating returns and competitive advantages. Reputation assessment companies offer services that can help other companies make market choices. In this project a system will be presented that can assess the reputation and reliability of a company quickly and easily. Due to the infinite amount of data that can be found online it will in fact be possible, through OSINT tools and payable services, to find information about a target and evaluate them according to the principles of corporate reputation.

Contents

1	Introduction	11
1.1	Motivation	12
1.2	Objective	12
1.3	Methodology	13
1.4	Outline of the Thesis	13
2	Literature review	15
2.1	Methodology	15
2.2	Research Questions	16
2.3	Corporate Reputation	16
2.3.1	Definition and key components of Corporate Reputation	16
2.4	Due Diligence	20
2.4.1	Concept and Definition of Due Diligence	20
2.4.2	Due Diligence Process	20
2.4.3	Due Diligence Audits	22
2.5	Measuring Reputation	27
2.5.1	Reputation index	27
2.5.2	Relative weights for Index Components	31
3	Project Overview	33
3.1	Systems Requirements	34
3.2	Stakeholders Survey	36
4	Information gathering Tools and Services	39
4.1	Information Management between Countries	39
4.2	Open Source INTelligence	41
4.2.1	Recon-ng	42
4.2.2	theHarvester	44
4.3	Payable Services	46
4.3.1	Cribis D&B	46
4.3.2	SocialNet	47
4.4	Company and Third Party interaction	48
5	Reputation System	49
5.1	System Architecture	49

5.2	Reputation Index Model	52
5.3	OSINT tools modules Implementation	56
5.4	Payable service interactions	60
5.5	Recursive phase and Report generation	64
5.6	User Interface	66
6	Reputation Evaluation Algorithm	69
6.1	Measures of Reputation Index Model	69
6.2	Assigning Values to the Index Model Categories	72
6.3	Algorithm Implementation on the System	73
7	Conclusions	77
7.1	Future Development	79
7.2	The Alfa Group Company	80
A	OSINT tools output Pictures	81
B	User Interface Pictures	85
C	Report Pictures	91

List of Figures

2.1	Due diligence Audit Requirements	22
4.1	Recon-ng console loaded	42
4.2	Recon-ng module options	43
4.3	Recon-ng “findsubdomains” module output	44
4.4	List of commands available on theHarvester	45
4.5	theHarvester, email addresses found	46
4.6	theHarvester, hostnames found	46
4.7	Cribis logo	47
5.1	System Architecture	50
5.2	Process Diagram	51
5.3	Cribis Request / Response model	61
5.4	Cribis XML Schema example	63
5.5	Network of people related to the Target Company	65
A.1	Email addresses retrieved by OSINT tools implemented on the system	82
A.2	List of hostnames retrieved by OSINT tools (first)	83
A.3	List of hostnames retrieved by OSINT tools (second)	84
B.1	User Interface home page	86
B.2	User Interface company section	87
B.3	User Interface people network section	88
B.4	User Interface report section	89
C.1	Example Report (first page)	92
C.2	Example Report (second page)	93

List of Tables

2.1	Definitions of Corporate Reputation	17
2.2	Due Diligence Dimensions and Environments	21
2.3	Due Diligence Auditing Process	23
2.4	Components of the Reputation Index	30
2.5	Relative Weights for Index Components in Corporate Reputation Model	31
3.1	Functional Requirement N.1	34
3.2	Functional Requirement N.2	34
3.3	Functional Requirement N.3	34
3.4	Functional Requirement N.4	35
3.5	Functional Requirement N.5	35
3.6	Functional Requirement N.6	35
5.1	The Reputation Index Model	52
6.1	Measures of the Reputation Index Model	71
6.2	Weights for Categories in Reputation Index Model	72

Listings

5.1	theHarvester Python Google module	57
5.2	theHarvester Python Parser Class	58
5.3	theHarvester Java Class	59
5.4	Cribis Business Company Report Method	64

1. Introduction

Today there is a lot of talk about reputation but it is a concept that began a long time ago in the field of business. It dates back to the pacts signed with a handshake, a gesture worth more than a thousand bills or notarial deeds. Of course, one cannot return to that gesture, nor can one live from nostalgia. Times have changed and today reputation is one of the many intangible assets of a company, representing the black gold of the third millennium.

The importance of this concept is confirmed through the numerous sources, the vast amount of people who talk about it and all the advantages that having a good reputation brings. In fact, enjoying a good reputation brings many advantages such as: increasing sales and standing out among competitors which consequently improve performance, especially from an economic point of view. Enjoying a good reputation generally influences the choices of customers, who tend to prefer one company to others, and also to business partners. Furthermore, the company's reputation is also fundamental for stakeholders, who take greater advantage in investing resources in a stable and high quality company. It's clear to see that reputation indicates the quality of the company without being a fixed indicator. In fact, this score changes positively or negatively as a result of various factors. The possibility of corporate reputation becoming negative results in a decline in performance and financial state. This risk is always around the corner and some factors can be: accidents that compromise the health and safety of workers, leaks, cyber attacks, negative balance sheets or simply disputes with the state.

The reputation of a company is of vital importance, as it can determine a growth or a decline in the business depending on whether it is positive or negative. From all these reasons, it is clear that the reputation must be constantly cultivated and reaffirmed. Citing a now famous quote from the famous financier Warren Buffet: "It takes twenty years to build a reputation and five minutes to ruin it. If you think about this, you will do things differently".

Having said that, how does one go about achieving a good reputation and how is the reputation score maintained? These issues are being addressed today and are of general interest for all sectors of business but all of them can be summarised into a single question. What is reputation and what are the components that define it? It is possible to find an answer to this question through numerous websites, papers and magazines thanks to the countless people who have tried their hand at this puzzle. From managers of companies, to financiers, to researchers; each of them has arrived at a solution with more or less the same results.

Given the importance of reputation, many companies today offer services that aim to improve corporate reputation, guaranteeing benefits to the customer in many respects. Other companies, on the other hand, offer a service that aims to evaluate the reputation of a company in order to identify its adequacy as a candidate for a business operation. The ability to evaluate a company is increasingly in demand by the market, in fact,

finding the right business partners or the right suppliers is of fundamental interest for every company.

The project that will be discussed throughout the following thesis will focus on defining the concept of reputation while identifying all the components connected to it in order to assess the reputation of a company. An automated service will then be created, with the ability to find specific information online regarding the target company in order to evaluate its reliability as a candidate for a business operation.

1.1 Motivation

The steady increase in the importance of reputation until today has not gone unnoticed, which is why many companies have begun to be concerned about their reputation. At the same time, investigative companies and services have been set up to assess the reputation of a company. These services are now widely used because of the assistance they provide when it comes to making market decisions.

Thanks to the agreement made between the University of Camerino and the Alfa Group Company of Porto San Giorgio, I was offered the opportunity to design and implement a system capable of assessing the reputation of a target company.

Given the need for the Alfa Group Company to carry out this work and my willingness to carry out the thesis in the company, I decided to accept the proposal. The proposed topic seemed immediately very interesting, because the concept of corporate reputation was addressed within OSINT. An environment of which I had already heard a lot and in which I had the opportunity to interact in the previous years of university studies.

The proposed project aims to create a system that evaluates the reputation and reliability of a company or a person with a VAT code, quickly and easily. The motive behind this project is simple. It is created to be used by any company that needs to evaluate the position of a target company as a candidate for a business action. There are many sectors into which it is thought that the system could be integrated but all share the same four business actions for which they might be searching a candidate. Companies could be seeking candidates for a Merger and Acquisition (M&A) or Joint Venture (JV) as well as simply suppliers or customers. All of these actions require candidates to be evaluated before a decision can be made, therefore this system will be of great use.

1.2 Objective

The main goal of the project is to develop a system which can assess the reliability of a target company. Therefore, the system will have to generate a report containing all the information retrieved about a target and moreover evaluate it before giving it a reputational score. The report generated by the system will be of fundamental importance to a company or an analyst evaluating the position of the target as a candidate for one of the business actions mentioned before.

The system must provide a user interface in which it is possible to insert a company name, VAT number and website URL. Once a request has been made in the user interface, the system functionalities will be executed in order to generate a report and to launch the evaluation algorithm. To achieve the main goal, the system has to perform several tasks. The main one is to make requests in different data sources in order to

retrieve information about the target inserted into the user interface. Therefore some OSINT tools and payable services will be used.

1.3 Methodology

In this section the methodology used to carry out this project will be described, who inspired the chosen methods and how the project was structured in order to achieve the goal of the thesis.

According to the Alfa Group Company and in agreement with the University of Camerino the project will be developed through the waterfall model.

“The waterfall model is a sequential design process, often used in software development processes, in which progress is seen as flowing steadily downwards (like a waterfall) through the phases.” [1]

The waterfall model is ideal for the development of this project because it defines the sequential phases of the project in which each phase depends on the results of the previous phase. In fact, it has a less iterative and more flexible approach making it ideal for software development. Through the phases of analysis, design, development, testing and implementation, the model will guide us towards the creation of the project.

The first phase of the project, namely that of analysis, was devoted to the review of the literature concerning corporate reputation. Given the importance attributed to reputation, a study and analysis of the literature is mandatory in order to understand the concept of reputation, but above all to identify the key components needed to assess reputation. The literature review was then extended to concepts concerning the measurement of reputation and due diligence.

During the second phase, the means were identified, namely OSINT tools and payable services, with which to find certain information online. These tools and services have been studied and tested, and during the third phase they were implemented in a system able to execute them automatically. The third phase is in fact related to the development of the reputation system. In this phase all the functions of the system were developed and a reputation index model was created containing all the key components and the parameters involved in the reputation assessment.

In the fourth phase of the project a reputation assessment algorithm was defined and implemented, based on the reputation index model.

1.4 Outline of the Thesis

The section that follows outlines briefly the chapters in the thesis:

Chapter 1 introduces the research project, provides a background of how the thesis was dealt with and explains what will be covered during the entire project. This chapter defines also the main goal of the thesis and how the different phases of the project were carried out.

Chapter 2 reviews the corporate reputation literature and includes the literature relating to associated concepts, such as due diligence. Within the literature their key components are defined and analysed. Moreover, how reputation is measured through an index is shown.

Chapter 3 presents a project overview in which the functional requirements of the entire project are defined and described, from the reputation index model that will

be defined to the final system that will be created. Additionally an analysis of the stakeholders involved is carried out.

Chapter 4 introduces the information gathering part. Once a definition has been provided of how online data is evolving country by country, it is explained how, and from where, the system will retrieve information related to a target, using the OSINT tools and payable services.

Chapter 5 explains the entire architecture of the final system created and presents all the different parts into which the system is divided. First, the reputation index model created through the literature review is defined, then the retrieving data phase is described using the tools and services shown in the previous chapter. Moreover the user interface of the system is described in detail.

Chapter 6 explains the evaluation algorithm created in order to assess the reputation through the key components of the reputation index model. It explains the different weights given to the key components based on the stakeholders involved and analyses the algorithm implementation of the system.

In chapter 7 the results and their implications in the light of an analysis of the reputation literature is discussed. An overview of the study is provided, followed by a summary and a conclusion of the entire project. This final chapter is rounded off with a discussion about future development and its implications for further research.

2. Literature review

What corporate reputation is, and how best to measure it, have become central business questions. The marketing literature and practitioner experience attest the importance of corporate reputation as an intangible asset that contributes to the market value of a company, to the need for more effective definition and measurement. To manage reputation effectively it must be defined and measured however there is a lack of agreement on its definition and on its key components; on the role of these key components; and on how to measure them.

This chapter is focused on the clarification, definition and understanding of the meaning of corporate reputation from an objective point of view, in particular a deep overview of the due diligence concept is covered to identify all the key components of the reputation. Moreover a research of how the reputation is measured through the key components is carried out in this chapter. All the management issues and all the process necessary to improve the corporate reputation, while certainly relevant, are outside the scope of this chapter.

2.1 Methodology

A research design will be implemented to explore the corporate reputation construct, its key components and a measurement approach in order to assess it. The first phase of research consists of text analysis of the academic literature on corporate reputation and its key components within a content analysis methodology. The second phase consists of research on similar topics and key components related to corporate reputation in order to retrieve more information during the definition of its meaning. The third phase, instead, includes research on how to assess the corporate reputation through the key components and parameters found during the previous phases.

The objective of the first phase is to gain a better understanding of the structure and the meaning of corporate reputation and other key concepts. After completing text analysis of the corporate reputation literature, it's necessary to identify the components of corporate reputation for each stakeholder segment, to explore the estimation of value and, moreover, to explore the meaning of the corporate reputation construct and its components.

In the second phase, a research on the due diligence concept is employed, not only in order to understand the meaning of due diligence and its process, but also because all the audits of the due diligence help to identify, from the various perspectives, the key components and parameters of reputation.

In the third, the focus is on evaluation of the reputation and the importance, to reputation assessment, of the various key components of corporate reputation. A study of some evaluation algorithms and how they work, will be carried out in this phase.

2.2 Research Questions

The research questions defined have been established in order to fully understand the meaning of corporate reputation and to discover the state of the art in how to assess the reputation. The following research questions, whose origins will be discussed as this thesis progresses, are discussed and tested in this study:

- **RQ1** *What is the reputation and why is it necessary to define it?*
This question aims at clarifying the meaning of reputation from an objective point of view, explaining also why it is so important to define it.
- **RQ2** *What is corporate reputation and which are its key components?*
This question aims at clarifying the corporate reputation and all the key components and parameters necessary to understand the concept.
- **RQ3** *What is due diligence and which are its key components?*
This question aims at understanding the due diligence concept, the due diligence process and also all the audit covered by the due diligence.
- **RQ4** *How can the reputation be assessed through the defined key components?*
This question aims at clarifying how to create an algorithm to evaluate the reputation.

2.3 Corporate Reputation

Corporate reputation has attracted interest from a wide range of academic disciplines. It is also a growing focus for business and media attention. This chapter examines the construct of corporate reputation, first by untangling the terminological problems that have been caused by the interdisciplinary nature of much of the earlier work in the area. Then, a research of how the reputation, that has been measured, is carried out. The chapter uncovers considerable confusion about how the corporate reputation is measured and how many companies tried to give their own measurement. An evaluation of several competing measurement models is also studied.

2.3.1 Definition and key components of Corporate Reputation

The identification of the definition of corporate reputation as a fundamental problem in the literature, and the appearance of a huge amount of articles discussing the definition, demonstrate the ongoing need for definitional consensus. Table 2.1 according to [15] presents definitions in chronological order given by some authors. Most of the articles examined used the Fombrun's definition.

“A perceptual representation of a company's past actions and future prospects that describes the firm's overall appeal to all of its key constituents when compared with other leading rival”

Fombrun [16] defines reputation as the overall estimation of a firm by its stakeholders, which is expressed by the net affective reactions of customers, investors, employees, and the general public. However, it is not clear why only affective reactions are allowed, and

<i>Author(s), Year</i>	<i>Definition</i>
Weigelt and Camerer, 1988	“A set of attributes ascribed to a firm, inferred from the firm’s past actions” [2]
Fombrun and Shanley, 1990	“The outcome of a competitive process in which firms signal their key characteristics to constituents to maximize their social status” [3]
Fombrun, 1996	“perceptual representation of a company’s past actions and future prospects that describes the firm’s overall appeal to all of its key constituents when compared with other leading rivals” [4]
Fombrun and Van Riel, 1997	“A corporate reputation is a collective representation of a firm’s past actions and results that describes the firm’s ability to deliver valued outcomes to multiple stakeholders. It gauges a firm’s relative standing both internally with employees and externally with its stakeholders, in both its competitive and institutional environment” [5]
Cable and Graham, 2000	“A public’s affective evaluation of a firms’ name relative to other firms” [6]
Deephouse, 2000	“The evaluation of a firm by its stakeholders in terms of their affect, esteem, and knowledge” [7]
Bromley, 2001	“... a distribution of opinions (the overt expressions of a collective image) about a person or other entity, in a stakeholder or interest group” [8]
Mahon, 2002	“a reckoning, an estimation, from the Latin reputatus – to reckon, to count over. The estimation in which a person, thing, or action is held by others . . . whether favorable or unfavorable” [9]
Whetten and Mackey, 2002	“Organizational reputation is a particular type of feedback, received by an organization from its stakeholders, concerning the credibility of the organization’s identity claims” [10]
Rindova, 2005	“Stakeholders’ perceptions about an organization’s ability to create value relative to competitors” [11]
Rhee and Haunschild, 2006	“The consumer’s subjective evaluation of the perceived quality of the producer” [12]
Carter, 2006	“A set of key characteristics attributed to a firm by various stakeholders” [13]
Barnett, 2006	“Observer’s collective judgments of a corporation based on assessments of the financial, social, and environmental impacts attributed to the corporate over time” [14]

Table 2.1: Definitions of Corporate Reputation

why cognitive components are excluded. Alternatively, Gray/Ballmer [17] define corporate reputation as a valuation of a company's attributes, performed by the stakeholders, what would almost completely exclude affective components. Hall [18] combines affective and cognitive components by formulating that a company's reputation consists of the knowledge and the emotions held by individuals.

According to [19], unlike scientific publications, more pragmatic literature avoids explicit definitions, particularly since the mere translation clarifies the content of reputation, and the only thing to discuss is how to operationalize it. Hence, at best, such publications state a set of attributes or categories to measure reputation. Nerb has screened the web pages of numerous PR and advertising agencies, and concludes that only few agencies explain what is meant by reputation, even if services titled "reputation management" belong to the standard program of many communication service providers.

The combination of affective and cognitive components points up that we conceptualize reputation as an attitudinal construct, where attitude denotes subjective, emotional, and cognitive based mindsets. 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. [19]

The proliferation of different methods measuring corporate reputation through different key components has raised the question of whether or not a standard can be established [20]. However, it seems there is no consensus at the moment. 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. Then a lot of companies tried to measure the corporate reputation from the German Manager Magazin to Fombrun and Harris with their reputation quotient (RQ).

In particular the Harris-Fombrun Reputation QuotientSM (RQ) [21] was developed by Harris Interactive, Charles Fombrun and Cees van Riel. The questionnaire they use to measure reputation consists of 20 items divided into six key components:

- Emotional Appeal (Have a good feeling about the company. Admire and respect the company. Trust the company a great deal).
- Products Services (Stands behind its products and services. Develops innovative products and services. Offers high quality products and services. Offers products and services that are a good value for the money).
- Financial Performance (Has a strong record of profitability. Looks like a low risk investment. Looks like a company with strong prospects for future growth. Tends to out-perform its competitors).
- Vision Leadership (Has excellent leadership. Has a clear vision for its future. Recognizes and takes advantage of market opportunities).
- Workplace Environment (Is well-managed. Looks like a good company to work for. Looks like a company that would have good employees).
- Social Responsibility (Supports good causes. Is an environmentally responsible company. Maintains high standards in the way it treats people).

The items were generated mainly from eight existing media rankings, including the Fortune AMAC, survey with some additional items from the image and reputation literatures. In considering how to measure corporate reputation, Fombrun and Shanley (1990) rightly argue against the use of single items in the Fortune scale to measure reputational attributes. Therefore they compute an index of overall corporate reputation derived from the eight attributes in the Fortune survey like showed in the previous chapter.

Fryxell and Wang (1994), in according to [22] criticize the Fortune index and point out that four of the eight items in the index refer to performance while constructs like innovation, corporate social responsibility and management quality are being measured by single items. Basically the Fortune index has been long discussed. In Fortune's survey, each company is rated relative to its leading competitors on eight characteristics using an 11-point scale (0 = poor, 10 = excellent) on the following eight attributes: Long-term investment value, Financial soundness, Wise use of corporate assets, Quality of management, Quality of products or services, Innovativeness, Ability to attract, develop, and keep talented people, and Community and environmental responsibility.

Even if this index, with its eight components, could be considered as untruthful or insufficient by somebody, it is useful for the research of this thesis, helping to better understand the corporate reputation concept. Back to the RQ-Index, it is the most advanced measurement model with respect to the others. A thorough discussion on validity and reliability of all described rankings cannot be given during this literature review since operationalizations are not given (Fortune, Fombrun) or not existent (ManagerMagazin).

Summary

An entire overview of the corporate reputation, its key components and some measured models are analyzed in this chapter. It became apparent that reputation is a corporation's most important competitive asset and notes that reputation can enable a stronger ability to attract and retain good employees, better margins, more attractive partners for mergers and acquisitions and more customers. This implies a positive effect of reputation on partnering in the context of mergers and acquisitions.

Regarding the reputation in M&A, the next chapter is going to show the reputation from the due diligence concept.

2.4 Due Diligence

Since corporate reputation can be applied in various fields, its use in ambitious M&A has become increasingly frequent. Specifically, we can talk about due diligence, the reputation field that deals with assessing risks and assessing the reputation from a legal and financial point of view. This chapter shows the due diligence concept and its key components needed to define the reputation. The meaning of the due diligence, the areas covered by the due diligence and why it's necessary conduct a due diligence process are themes analyzed in this chapter.

2.4.1 Concept and Definition of Due Diligence

There is no dictionary definition of due diligence. There is no standard legal definition either. A lawyer would probably define it as follow:

“The investigation of the assets and liabilities of a company or business for the purposes of buying or selling its assets”

According to [23], any search of a law dictionary reveals the meaning of ‘due’ as payable or immediately enforceable. ‘Diligence’ involves care, attention and application. Together, therefore, the words have an interesting emphasis on the enforceability of the process. Having regard to the purpose of due diligence and the evolution of the term, it is no surprise that the legal process has become increasingly comprehensive as regulatory and other business frameworks have developed.

Due diligence activities, as with everything, need a starting point. If there are talks about a possible merger, each party to the transaction must be willing to commence a due diligence activity. However, this is where the definition of the activity of due diligence can become blurred. In a merger situation, there would normally be a significant amount of due diligence prior to any informal or formal conversation. The diligence required is to determine whether there is enough information that can lead to conversations about a possible merger. Thus the starting point for any due diligence activity is never one single step with a single starting point. Consequently, the actions surrounding due diligence must be adaptable within a framework that places the organization and its owners, employees and advisers in a constant state of data collection and data organization that can support whatever process is being started.

2.4.2 Due Diligence Process

The due diligence process refers to the review of a target company to make sure that the purchase does not pose any unnecessary risk to the acquirers shareholder. By inquiring into all relevant aspects of the past, present, and predictable future of the target firm, one clarifies benefits and liabilities in a proposed acquisition. [24]

Table 2.2 illustrates the internal and external environments and examples of assets that need to be assessed during a due diligence process. The focus of a traditional due diligence process has been the tangible, internal environment. Financial and legal experts audit the “hard assets” and attempt to determine potential liabilities or future projected growth scenarios after the company is acquired. The attention of the auditors, e.g., lawyers and accountants, is primarily focused on verifying historical data and affixing value to the tangible assets of the company.

Internal Environment	External Environment
<p>Tangible</p> <ul style="list-style-type: none"> • cash • plant equipment • accounts receivable • patents/trademarks • technology • inventory 	<p>Tangible</p> <ul style="list-style-type: none"> • share of market • supplier/distributor contracts • physical location
<p>Intangible</p> <ul style="list-style-type: none"> • quality of leadership • training of personnel • corporate culture • quality of infor/analysis operating systems • loyalty of personnel • trade secrets • data bases • personal/professional networks 	<p>Intangible</p> <ul style="list-style-type: none"> • brand product awareness • customer loyalty • competitive positioning

Table 2.2: Due Diligence Dimensions and Environments

The information, which has not been historically included in the due diligence process, is normally considered to be intangible, “soft” assets of the company. These internal soft assets, i.e., quality of management, personnel, corporate culture and customer loyalty, are essential elements in the future success or failure of the acquisition.

According to [24], the intangible assets of a company can be classified into six categories: the “having” capabilities, e.g., products, distribution, reputations; the “doing” capabilities, i.e., knowing how to do something, leadership; “people dependency” e.g., reputation; “people independency” e.g., databases; protection by law, e.g., trademarks; and no protection by law, e.g., organizational networks. It would be difficult to argue that these intangible dimensions of a company would not impact the attractiveness, purchase price, or future value to the acquiring company. There are even significant tax issues associated with these intangible assets. [24]

The scope of due diligence during the 1990s needed to be defined to identify domains of information that are required to make a more informed acquisition. The picture 2.1 shows seven fields that need to be investigated in a comprehensive due diligence process.

Those involved in more traditional due diligence would argue that the fields identified in the picture 2.1 have been encompassed within their due diligence in the past. The basic organizational matters, e.g., bylaws, corporate charter, governance structure, are included in a traditional due diligence process. The management of the company to be acquired is examined with regard to executive remuneration, employment contracts, arrangements for future consideration (buyouts), and other tangible aspects relative to the management of the company.

Importantly, however, the traditional due diligence process primarily focuses on tangible assets or documents relative to the operation of the company. What is being

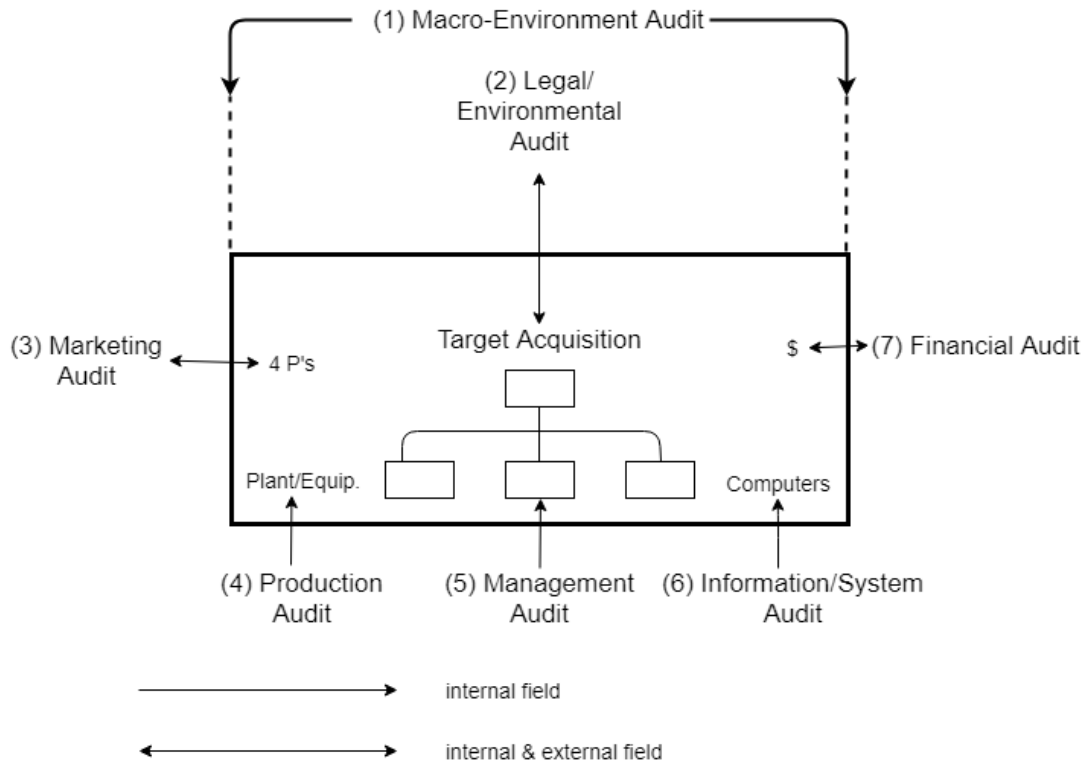


Figure 2.1: Due diligence Audit Requirements

recommended is that each of the seven fields identified in the picture 2.1 be audited, examining both tangible and intangible dimensions of the function.

2.4.3 Due Diligence Audits

According to [24] the table 2.3 outlines the types of audits that need to be conducted during due diligence to adequately examine potential acquisition candidates. Moreover it also details the sequence of audits and who has the primary responsibility for conducting each of the due diligence audits. The following discussion of audit procedures, which can be used in each of the seven audit fields, illustrates procedures that should generate adequate information on tangible and intangible dimensions for each area that needs pre-acquisition investigation by the acquisition management team.

Macro environment Audit

There is a wide variety of techniques and methods used to assess and monitor the macro-environment for evidence of change. This evaluation of the “climate” of business becomes critical in the acquisition of a company, particularly if the company to be acquired is not in the same industry as the acquiring company. [25], [26] Discovering “signals” that may precede significant changes in the macro-environment before they have economic, social, or strategic impact is the goal of environmental scanning. Below there is a list of areas analyzed during an environmental scanning:

- Impact of the macro-environment on industry: Assessment of political, economic, social and, to some degree, psychological trends that influence behaviour in the

Type of Audit	Characteristic			
	Sequences	Time	Source of Data	Outcomes
1. Macro environment	3	M	External	Industry history/trends
2. Legal/Environmental	2	M	External/ Internal	Organizational documentation, contingent liabilities, existing/potential internal and external hazards
3. Marketing	4	L	External/ internal	Product positioning, consumer loyalty, sales effectiveness
4. Production	4	M	Internal	Efficiency of production process, capacity, equipment assessment, synergy
5. Management	4	L	Internal	Quality of managers, organizational structure, succession, training, team, corporate culture
6. Information system	5	S	Internal	Hardware/ Software compatibility
7. Financial	1	L	Internal	Historical/ projection scenarios/ tax issues

Table 2.3: Due Diligence Auditing Process

target industry.

- Competitor profiles: Identification of competitors and their strategic positioning in the marketplace, relative competitive advantage, strengths and weaknesses.
- Environmental opportunities and risks: Potential growth opportunities as well as risks inherent in the macro-environment that could reduce the success of the company.

Legal / Environmental Audit

Traditional legal due diligence is conducted by outside legal counsel to the acquiring company. Briefly the functions that are typically included in a legal audit are: basic organizational matters; ownership of securities; banks and borrowing; financial history; litigation; general regulatory data; real property; personal property; intellectual property rights; contractual management issues; labor contracts and history; and insurance. These activities constitute the core of a traditional legal audit of a potential acquisition. One additional area that has received increased attention from the counsel to the acquiring company has been the environmental audit. The objective of a comprehensive environmental audit is to ascertain the degree or status of a company's compliance with all statutory and regulatory authorities to which it is subject. An environmental audit incorporated the following elements:

- materials entering the facility, such as hazardous material or solvents.
- materials, personnel, activities, and conditions occurring on-site, such as workers' and the public's right-to-know, processes or spills of fuels or other materials.
- effects on site attributes, such as wetlands or zoning considerations.
- emissions and workers leaving the site, such as leaking underground fuel tanks, air emissions, or wastewaters. [27]

Marketing Audit

The purpose of the marketing audit is to determine how marketing functions are being performed, an appraisal of the execution of these functions, and recommendations of how to improve the marketing effort. To effectively assess the marketing efforts of the target company, the following areas must be investigated:

- Marketing environment review: aggregate demand for the product, specific market segments appealed to by the company, major shifts/trends in each segment, and forecasted trends for the segments. Further, this stage of the marketing audit examines customer motivation, image of the product/service produced by the target acquisition, level of satisfaction, and/or brand loyalty to the company's product/service.
- Marketing system review: this stage of the marketing audit examines the marketing objectives and the marketing programs. This assessment of marketing programs should include their activities, procedures, and personnel.
- Marketing functions review: this phase of the marketing audit focuses attention on each of the marketing functions, i.e., product development, pricing, promotion, physical distribution, etc. [28]

The most difficult aspect of the marketing audit is making definitive assessments of intangible aspects of the company's performance. These include: level of brand loyalty by customer; consumers' attitudes toward the company and its products; relative market position among competitors by market segment; qualitative assessment of advertising and promotion effects; aggregate value of market position, brand awareness/loyalty/distributor quality/loyalty and, in general, the logo/trademark/brand value. [18]

Production Audit

Acquisitions that include a manufactured product require an evaluation of the manufacturing process and the production capabilities of the intended acquisition. This audit process can be divided into three related areas:

- Physical plant and equipment: the physical characteristics and dimensions of the physical facilities need to be determined.
- Manufacturing systems assessment: the manufacturing process must be examined to determine the manufacturing task, utilization of facilities, equipment, and personnel needed for various combinations of products being produced.

- Personnel: the manufacturing personnel, including the management, supervisors, and production workers, should be interviewed during the due diligence process.

Management Audit

The legal audit examines personnel issues but they tend to be the formal, technical issues associated with the company's personnel. Additional insights into the management of the company need to be undertaken during due diligence. The acquisition may have an adverse effect on the company's existing employees, and these employees may have a significant impact on the success of the company after acquisition. [29] The key dimensions of a management audit are as follows:

- Organization structure: an analysis of how the company is formally organized and problems associated with the organizational structure.
- Personnel assessment: an in-depth evaluation of key executives and managers to determine their strengths and weaknesses as well as their predisposition toward the potential acquisition.
- Compensation/benefit program: the means for remuneration of employees in the past has to be examined to determine compatibility with the acquiring company.
- Management infrastructure: an assessment of the planning system, management control system, and management training and development programs is needed.

Information System Audit

One area that has become more important to the successful merger of the parent and acquired company is their information systems. Both the hardware and software aspects of the systems are important dimensions of the acquisition that need to be explored in due diligence. An audit of the information system should encompass three integrated issues: technology issues, management issues, and merging/transition issues.

- Technology issues: Compatibility of hardware with the acquiring company, as well as hardware reliability, capacity, and use and maintenance records are vital information. Beyond the issue of hardware architecture, other technology issues such as vendor agreements, i.e., maintenance, networks, facilities, are important. The type of network used, location of terminals and microcomputers that permit network access by internal users, customers, vendors, and suppliers must also be determined. Software is also considered to be a technology issue, and one that can be costly if not compatible with that of the acquiring company.
- Management issues: the human resources information system audit procedure focuses on the number of personnel and the direct cost associated with maintaining the information system.
- Merger/transition issues: This declaration will lead to an assessment of personnel, hardware capabilities and, ultimately, provide the templates for an integration plan for the information system function.

Financial Audit

The basic financial audit provides the platform for the acquisition team to project future growth/profit scenarios for the company. It also encapsulates the position of the company on an accrued basis of the following: existence or occurrence of material events and/or matters; completeness of the data without selective omission of dates; rights and obligations of the assets of the company; valuation and/or allocation of assets and liabilities; and disclosure of potential liabilities that have not been determined at the time of the audit.

Summary

The due diligence process conducted during mergers and acquisition should provide decision-makers with information on opportunities as well as potential problems. Often, due diligence has been viewed as a mechanical “verification” of legal, accounting, and tax matters, while all the others intangible functions have received cursory examination by members of the due diligence team. In this chapter instead, was presented, beyond the traditional legal, accounting/financial, and environmental audits, all the other audits that a due diligence process should cover, such as macro-environment, production, management, marketing, and information system audits were presented.

2.5 Measuring Reputation

The previous chapters showed the reputation concept from a corporate reputation and due diligence point of view. Underlining the main key components of each one, it was possible to see that both of them share the same categories, in which corporate reputation is more focused on products, services and stakeholders in general, while due diligence is more focused on the legal, financial and tax related matters.

Once understanding the reputation concept, the literature review has been focused on how to measure the reputation. This chapter presents a model index and weights for the index components that will help to understand how a measurement algorithm reputation should be developed.

According to the research carried out into the previous chapter and according to [30], a comprehensive desk research shows that existing reputation measurement tools are based on the following categories, which can be described as first or second order formed attributes:

- Financial performance
- Quality of products and services
- Quality of employees
- Quality of management
- Market leadership
- Customer orientation
- Attractiveness
- Social responsibility
- Ethical Behaviour
- Reliability

Knowing that the main attributes or categories necessary to measure the reputation are those above, a reputation index can be presented, as well as, the weight of the key components of the index.

2.5.1 Reputation index

The paper taken into consideration, [30], suggests a corporate reputation index based on specific measures relating to: products, employees, external relationships, innovation and value creation, financial strength and viability, strategy, culture, and intangible liabilities. The main goal of the paper is to provide a comprehensive set of components for the index and an initial set of illustrative measures from which to begin empirical research. While the components of corporate reputation are comprehensive, the specific component measures featured in the index will develop over time as research progresses. The authors of this papers think that the reputation index as a standardized set of common as well as unique component measures that would be consistent across companies and industries.

According to [30], Table 2.4 illustrates the components of the corporate reputation index and includes illustrative measures for each of the components. In contrast to other measures of intangible assets, the corporate reputation index is based primarily on internal assessment of non-quantitative factors and does not rely on market values or asset values. Instead, the components of corporate reputation are evaluated on the basis of internal and external information. The basic index includes (common) factors that should be relevant in assessing corporate reputation in general.

The develop of the reputation index was carried out in different key groups who interface with the organization to assess their opinions of corporate reputation, from customers to suppliers, employees, partners in alliances or partnerships, and even from competitors.

Take a look at Table 2.4, the Products / Services component provides the essential interface with the customer and a major driver of corporate reputation. The index must assess the extent to which the reputation of a store's name affects the overall corporate reputation. In addition to consumer awareness, the index must measure the strength of the corporate reputation in terms of the attributes of the products or services offered, including quality. Similarly, the extent of external quality failures as evidenced by warranty and liability claims is also a part of reputation.

The employees are the means by which a corporate reputation is created. The index could evaluate the length of time that employees remain at the company and the extent to which new applicants seek jobs at the company. Training and development reflects an investment in both the employees (by the firm) and in the company (by the employees). The personal reputation of the CEO should be evaluated. Similarly, the competency and turnover of all upper management should be assessed as well. Management must be competent to make decisions and their motives should also be examined.

External relationships, in addition to the relationship with customers, are important components of corporate reputation. Key relationships with suppliers, partners, investors and even competitors have contributed to corporate reputation. The reputation index should measure the quality of the suppliers and even the suppliers' suppliers. Similarly, the corporate reputation of any strategic alliance or joint venture partner can have an impact on reputation. The extent to which investors have confidence in the corporation is an informative measure that can be determined through analyst activity, share volume, and surveys. It is also necessary to examine corporate reputation from the perspective of competitors.

Both innovation and value creation are essential corporate attributes in the new economy, so some measures to evaluate the reputation have been taken into consideration. Measures on this component should be centred on customer-focused attributes. For example, growth relative to customer needs, new product/service development and customer retention are useful measures.

The financial component is always under the sights of all, for this reason, as seen in the previous chapter, there is a great discussion about it. in this index the financial component has not been taken very much into consideration. The extent to which users of financial statement information can trust the information affects is the aspect covered.

The strategy of the company should be at the focal point of decision-making. Thus, the extent to which strategic priorities address corporate reputation issues will emphasize the importance of reputation. The reputation index should assess strategic priorities and consider how these priorities are integrated across business units. The

Index components	Illustrative measures	Anchor scale values:	
		least desirable	optimal (ideal)
Products/services	Quality associations	Almost none	Highest
	Public awareness of corporate name and products/services	Almost none	Highest
	Extent of brands and umbrella brands	Single brand item	Numerous brand lines
	Warranty claims	Often, numerous	Never
	Liability claims	Often, numerous	Never
Employees: All levels	Employee satisfaction with employer	Almost none	Highest
	Turnover	Common, extensive	Almost none
	Exit interviews	None conducted	Formal, informative
	Number of applicants for open positions	None, unfilled positions	Excessive, high interest
	Training and development efforts	None or rare	Extensive
	Employee feedback relative to meeting employee needs	None or rare	Highest
	Coordination and communication efforts across functional and business areas	None, isolated, lack of information flow	Extensive and regular
	CEO personal reputation only	Almost none	Highest
	Competency	Poor	Highest
	Turnover	Common, extensive	Almost none
Upper management	Compensation and evaluation packages and goal congruence with strategic objectives	Incongruent, at odds with long-term objectives	Congruent and contributes to achieving long-term objectives
	Information collection from subordinates	None, isolated	Regular, participative
External relationships Suppliers	Major supplier quality	Almost none	Highest
	Relationship quality of major suppliers	Poor, no level of trust	Highest level of trust
	Relationship duration for major suppliers	Beginning	Enduring, long-term
	Quality of suppliers for suppliers	Almost none	Highest
Partners	Existence of alliance relationships	None	Numerous
	Longevity of alliance relationships	None or beginning	Enduring, long-term
	Recognition of key strategic partners	Unknown	Well-known
	Reputation of key strategic partners	Poor	High
	Joint venture contractual agreements	None	Numerous

Index components	Illustrative measures	Anchor scale values:	
		least desirable	optimal (ideal)
Competitors	Industry participation	Isolated	Active, exchange of info
	Competitor response to key corporate initiatives	Ignores	Immediately matches or responds to actions
Investors	Market premium	None	Highest
	Market stability	None, unstable	Long-term stability
Environment	Environmental policy	None	Formal, well-developed
	Dedicated employee positions	None	Dept. and sr. manager
Innovation	Liability claims	Often, numerous	Never
	Regulatory intervention	Often	Never
	Formalized program to generate and evaluate innovation	None	Mature, successful
	Growth relative to customer needs	Stagnant (poor)	Steady and consistent
Value creation	New product/service development	None	Extensive at all stages
	Identification and responsiveness to customer needs	Unaware of customer needs	Anticipates and meets all needs
Financial strength	Customer retention	Frequent loss	No customer defection
	Information content of annual report	Almost none	Highest
Strategy	Additional disclosures	None	Numerous and extensive
	Strategic priorities relative to reputation	Ignores reputation	Highest priority
	Integration of strategy across business units	None	Complete formal and operational integration
Culture	Management control system fostering consistency	No formal system in place	Formal system with perfect consistency
	Ethics policy	None in Place	Highly effective
	Reporting procedure for ethics violations	None in Place	Highly effective
	Upper management attitudes	Unethical or ignores	Corporate priority
Intangible liabilities	Ethics committee on the board	No	Yes
	Inadequate research and development process	No formal process	Highly developed and successful process
	Lack of adequate information infrastructure	No infrastructure	Well-developed
	Organizational structure – lack of flexibility	Totally inflexible	Highly flexible
	Bad word-of-mouth among customers	Numerous and common	Non existent
	Inadequate distribution channels	Numerous and common	Non existent

Table 2.4: Components of the Reputation Index

lack of a strategy could lead to a loss in corporate reputation.

The corporate reputation evaluation process should involve specific attention to the ethical climate of the organization. Ethical violations have the potential to create significant negative reactions from all stakeholder groups. Thus, the index should take into account the existence and extent of a corporate ethics policy.

2.5.2 Relative weights for Index Components

According to [30], it suggests a nine-point scale to assess the magnitude of the measure. The scale values for the measures in Table 2.4 are anchored with 9 as an ideal or benchmark and 1 as the lowest, or least desirable measure on the scale. The responses to the scales for the individual measures should be averaged for each separate component of corporate reputation. Once there is an aggregate measure (value of 1–9) for each component, an overall measure can be created by applying weights to each of the components and summing the values.

Table 2.5 suggests a range for the component weights in the index, it is only a range of weights given by [30], so the weights should be changed according to the field in which the reputation is applied. In the following chapters the fields of our reputation system will be defined.

Index Components	Range of Weights (sum to 100%)
Products or services	30-60%
Employees / suppliers	1-20%
External relationships / alliances	1-0%
Innovation	0-20%
Value creation	5-20%
Financial strength and viability	0-10%
Strategy	1-10%
Culture	1-0%
Intangible liabilities	0-0%

Table 2.5: Relative Weights for Index Components in Corporate Reputation Model

[30], Suggesting weights to apply to the various components of corporate reputation, considered the effect of the product or service offered to be of primary importance. The range for the weight of this component might range from 30 to 60 per cent. Employees, external relationships, innovation, and value creation can each reach a maximum of 20 per cent. All of these components are important, yet the relative impact on corporate reputation depends upon the strategic mission and operational efforts of the company at a given point in the corporate life cycle. Thus, these components may vary significantly in importance according to specific company characteristics and priorities. Although the minimum value for these components ranges from zero to one, value creation has a minimum weight value of 5 per cent. This component is so critical in terms of the reputation of the company that there is a higher minimum value. The annual report, strategy, culture, and intangible liabilities components have a maximum weight of 10 per cent as some of the elements of these components are implicitly considered in other areas.

The final step in creating a corporate reputation index is to translate the overall single scale measure (range of 1–9) to a classification ranking. [30] employed a nine classification categories for the index that are associated with descriptions of various standards of corporate reputation. This ranking is similar to ratings for bonds as developed by Moody's.

Summary

The objective of this chapter was to provide an input from which to begin an investigation into the assessment and measurement of corporate reputation. Although it is just from a point of view of the paper analysed, it was useful for understanding how a measurement reputation algorithm should be developed for the thesis final system. It will be necessary to refine the weighting scales for the key index components but a good input was provided by this chapter. The creation of a reputation index is the first step in standardising measurement and management of the most central intangible asset of all, the Reputation.

3. Project Overview

A general idea of the project was described in the first chapter, where the main goal was defined as well as the target of the project. The concepts of corporate reputation and due diligence with their key components have been identified, defined and described through the literature review. Instead, this chapter aims to describe the whole project in a more detailed and technical way, starting from its description with all its development phases, to the analysis of the system requirements and the definition of the stakeholders considered in the thesis.

According to and in agreement with the Alfa Group Company, the “Reputation system” project intends to design and develop a system that assesses the reputation and the reliability of a company or a person in possession of a VAT code. In detail, the system should be able to gather specific information about the defined target and create a network of people connected to it or that have an agreement with it, in order to have an idea about the target.

The type of information to be searched will be first defined in an index model containing all the key components and parameters necessary to assess the reputation from an objective point of view. The data retrieved and the list of people connected to it will be used after, to give a reputational score to the target and to analyse this target from a business point of view. In fact, according to the Alfa Group Company, the target will be analysed through four possible candidates, in case the target is chosen for a Merge and Acquisition or in case of a Joint Venture or in the event that it is chosen as a possible Supplier or customer. The evaluation model is the same one as the model containing all the key components to assess the reputation and it will be the same for each candidate, with the difference that different weights will be applied to the key components during the calculation of the score by the algorithm. All the data retrieved, from the ones found through the defined model to the network of people connected to the target, will be stored in a report which will be generated as the output of the system after all the operations are done.

The system that will be developed during the whole project, as it has been possible to deduce, will have to perform different functions, from some more complex and long to others less complex and distinctly shorter. During this chapter we will define the main functions that the system will have to implement by going through each one in detail. Moreover, the areas in which the target is seen as a possible candidate will be a described and analysed.

3.1 Systems Requirements

In order to design and implement a “Reputation system”, according to the Alpha Group Company, the concept of reputation must be defined. Thanks to the company itself that provided the research keys, the first phase of the project focused on defining the reputation from an objective point of view. A study was carried out during the literature review phase, in which the keywords provided were described and defined, allowing us to fully understand the meaning of reputation today.

The first functional requirement of the project, as shown in the table 3.1, is therefore that of defining an index model containing all the components and parameters necessary for defining and assessing the reputation for which the whole part of literature review has been carried out. This model will be inserted in the system in order to organise the ways in which to make requests and the way to parse the information retrieved. Moreover, it will be used as a structure for the final report.

ID	RF1
Name	Model definition
Definition	Create a model containing the key components needed to define the reputation. The model will be divided into parameters and categories.
Motivation	The defined model will be used as a grid that establishes the parameters necessary to assess the reputation of a company.

Table 3.1: Functional Requirement N.1

ID	RF2
Name	Data gathering
Definition	Design a system that, given an initial target, makes requests on the main Search Engines and services in order to retrieve the necessary data
Motivation	The data found must be processed and included in a report

Table 3.2: Functional Requirement N.2

ID	RF3
Name	Information parsing
Definition	Design a system that analyses the data found in RF2 and returns only the information required in the parameters of the model defined in RF1
Motivation	The information returned will be used to create the final report and to give a reputation score of the target

Table 3.3: Functional Requirement N.3

Starting in chronological order of functions that the system should perform, the second functional requirement is shown in table 3.2, in which the gathering information phase should be carried out. In this phase the system, once a target is defined, will

ID	RF4
Name	Recursive research and analysis
Definition	The system will have to carry out the RF2 and RF3 recursively on each company and / or person found during the first RF2
Motivation	Recursive research and analysis is necessary to better understand the links that the target company has with other companies or people

Table 3.4: Functional Requirement N.4

ID	RF5
Name	Final report
Definition	The system will generate a final report for the target defined with the structure of the model chosen in RF1 and the information returned in RF3 with, moreover, all the links and the relative information from RF4
Motivation	The final report will be the output of the system, the one that the customer will receive and view

Table 3.5: Functional Requirement N.5

have to make requests on the main data sources through Search Engines and services in order to retrieve the necessary data.

The third functional requirement is the information parsing, table 3.3, in which once the data is retrieved through the requests made in the previous phase, it must be parsed using the defined index model. The information returned will be then used to create the final report and give a reputation score of the target.

Through the information retrieved during the previous phase, a list of people or companies connected to the target will be found. In this case the system will have to perform a recursive research and analysis of information about this list, as shown in the table 3.4, repeating the data gathering phase and the information parsing phase. This operation will be repeated more times in order to discover the lists of people or companies connected to the first list of people or companies which are connected to the target. The network of people in contact with it will also be useful in assessing the reputation of the target.

Table 3.5 instead defines the way in which the system should return the amount of information retrieved during all the previous phases. In fact, the output of the system will have to be a report, a document with the same structure of the model defined first, containing all the information organised by key components and parameters of the model.

ID	RF6
Name	Target evaluation
Definition	The system must evaluate the report created in RF5 with a score according to a previously defined algorithm
Motivation	The score given in the report will be an indicator to understand the degree of reliability of the defined target

Table 3.6: Functional Requirement N.6

The last functional requirement is the target evaluation as shown in the table 3.6. The system will have to evaluate the report created in the previous phase, giving it a score based on a defined scale and algorithm, showing the degree of reliability of the target.

3.2 Stakeholders Survey

The “Reputation System” project, created within the Alfa Group Company, is aimed at creating a report based on a reputational model in order to assess the reliability of a target. Since the target is a company or a person with a VAT code, the relationships that can be established with it will concern the business world, but more specifically, the target will be considered for a Merge and Acquisition action, for a Joint Venture, or as a possible supplier or customer. As agreed with the company itself, in fact, the possible actions that can be taken with the chosen target will fall on these four areas. A more detailed explanation of the four is provided below.

Mergers and acquisitions (M&A) is the area of corporate finances, management and strategy dealing with purchasing and/or joining with other companies. In a merger, two organizations join forces to become a new business, usually with a new name. Such transactions typically happen between two businesses that are typically of similar size and stature and which recognize advantages the other offers in terms of increasing sales, efficiencies, and capabilities. In an acquisition, on the other hand, one business buys a second and generally smaller company which may be absorbed into the parent organization or run as a subsidiary. The end result of both processes is the same, but the relationship between the two companies differs based on whether a merger or acquisition occurred.

Joint Venture (JV) is a business arrangement in which two or more parties agree to pool their resources for the purpose of accomplishing a specific task. This task can be a new project or any other business activity. In a joint venture, each of the participants is responsible for profits, losses, and costs associated with it. However, the venture is its own entity, separate from the participants’ other business interests.

A supplier is a person or a business entity that supplies goods and services. It is typically described as the figure that is paid for goods that are provided, rather than the manufacturer of the goods itself. It is possible for a supplier, however, to sometimes operate as both a seller of goods and a manufacturer. Large retail store chains, for example, generally have a list of vendors from which they purchase goods at wholesale prices that they then sell at retail prices to their customers. In addition, a supplier can act as a business-to-business (B2B) sales organization that provides parts of a product to another business to make an end product.

A customer is an individual or business that purchases another company’s goods or services. Most public-facing businesses compete with other companies to attract customers, either by aggressively advertising their products or by lowering prices, in an effort to expand their customer bases. The other part is Business customers, also known as industrial customers, who purchase products or services to use in the production of other products.

The target will then be evaluated as a candidate of each possible action described, in order to have a score for each of the areas and then know which action is best to take or not. The reputation index model containing the key components and the parameters of the reputation is unique, while the weights given to key components during the

evaluation algorithm in the Chapter 6, will change according to the four areas.

Summary

The “Reputation System” project was thus analysed specifically, describing and motivating all its functions. The four areas in which the target is considered a possible candidate have been described, thus shifting attention not only to the main objective of the project, which is the creation of the report, but also which action is best to take with the target itself. The algorithm that will be developed will in fact show different scores based on the areas described above, giving way to those who will analyse the system report to understand in which areas the target is more valid.

Thus identified the functional requirements of the system and the areas in which the target is considered as a possible candidate, in the next section we will define and describe the tools and services with which to find the information to create the report.

4. Information gathering Tools and Services

Internet now offers an incredible wealth of heterogeneous resources: news, articles, comments, videos, numerical data, images, digital books, audio recordings and many other things. The amount of information grows exponentially at an astonishing rate thanks to factors such as the increasingly widespread diffusion of the Internet, the progressive digitalisation of all cognitive resources and above all thanks to a system of information production radically decentralised. All you need is a computer, a tablet or a smartphone and connection to the Internet to become an information producer. Anyone can produce information on the network, react to information entered by others and propose facts, ideas or criticisms.

With the exponential growth of online information, many companies have created services or programs that can search for specific information from the Web. Similarly, other companies have created real databases, exploiting the immense amount of data generated, without taking into consideration the privacy of individuals. Many of these services, as well as the access to some databases, are free of charge.

Through this premise, this chapter will describe the main tools and services used to find the information needed to assess the reputation of a given target. Other topics discussed, are the differences in information management from continent to continent, and the use of third parties to find information not retrieved on the internet.

4.1 Information Management between Countries

The internet has made it easy to access information by visiting a website, or to buy goods and services at the touch of a button. But most consumers aren't always fully aware that in doing this, the organisations they deal with online are collecting vast amounts of personal data about them. The ease and sophistication of data collection means that thousands of companies not only collect personal details, but store it in often insecure locations, share it with third parties or move this data across borders to support their businesses [31].

With many security breaches now well publicised by the media, consumers are increasingly becoming aware about what happens to their data and have looming privacy concerns about what is being stored and processed, and by who. Every country or continent has recognised the lack or inadequate use of information privacy law or data protection laws, so they have updated their rules. Over eighty countries and independent territories around the world have now adopted comprehensive data protection laws. The European Union has the General Data Protection Regulation (GDPR), in force since May 25, 2018. The United States is notable for not having adopted a com-

prehensive information privacy law, but rather having adopted limited sectoral laws in some areas like health and credit. While in Asia and its countries, there aren't any specific rules or laws. China is still implementing its Cyber Security Law and the stepping up of data protection laws, instead in Japan and Australia there is an overall trend towards stricter enforcement and greater public awareness of their rights under data protection laws.

As said before, there is a big difference between the United States and Europe in terms of data protection. The GDPR is a regulation in EU law on data protection and privacy for all individual citizens of the European Union (EU) and the European Economic Area (EEA). The GDPR aims primarily to give control to individuals over their personal data and to simplify the regulatory environment for international business by unifying the regulation within the EU. Controllers of personal data must put in place appropriate technical and organisational measures to implement the data protection principles. Moreover, no personal data may be processed unless it is done under a lawful basis specified by the regulation, or unless the data controller or processor has received an unambiguous and individualized affirmation of consent from the data subject. The data subject has the right to revoke this consent at any time.

Instead, in the U.S., although partial regulations exist, there is no all-encompassing law regulating the acquisition, storage, or use of personal data. In general terms, whoever can be troubled to key in the data, is deemed to own the right to store and use it, even if the data was collected without permission, except to any extent regulated by laws and rules. For instance, the Health Insurance Portability and Accountability Act of 1996 (HIPAA), the Children's Online Privacy Protection Act of 1998 (COPPA), and the Fair and Accurate Credit Transactions Act of 2003 (FACTA). These are all examples of U.S. federal laws with provisions which tend to promote information flow efficiencies.

According to [32], all the talk about data privacy are caught up in political wrangling but the different approaches have practical consequences for people, too. For example in Europe people can ask search engines like Google and Microsoft's Bing to remove links to negative articles about themselves on European versions of those sites. However, in the U.S. there is no blanket ruling that allows people to delete or remove negative information about themselves online. Moreover, in Europe, people can ask any company to send you details about what data it holds on you and what that information is used for and in most of the cases, companies must hand over the files within a month. Instead in the U.S., there is no single federal law or standard people can rely on to obtain copies of their records. But there are industry-specific rules. Patients, for instance, may request copies of their medical records from health-care providers.

A different approach by each country to data protection laws leads to a different way of retrieving information on a company for each country. For example in the USA, where most information is public, it's possible to just use Google or some online free service to gather information. Instead, in Europe everything is more difficult. Some information might only be retrieved through payable services or, in most of the cases, it cannot be retrieved at all.

Through this chapter, a different approach has been given to the research information phase. In fact, knowing that the final system is going to search information of an Italian company or at least of a European company, it is not possible to use free services or tools to collect information that would have been used in the case of an American company. In the following chapters, the tools and services which have been used during the thesis, to search for information, will be explained.

4.2 Open Source INTelligence

OSINT (Open Source Intelligence) refers to all the information that is publicly available. There is no specific date on when the term OSINT was first proposed; however, a relative term has probably been used for hundreds of years to describe the act of gathering intelligence through exploiting publicly available resources. The U.S. Department of Defense (DoD) defines OSINT as follows:

“Open-source intelligence (OSINT) is an intelligence that is produced from publicly available information and is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement”

The history of OSINT is relatively recent. According to [33], It was introduced during World War II as an intelligence tool when the United States established the Foreign Broadcast Information Service (FBIS) to monitor publicly available information. The FBIS continued its work in exploiting OSINT sources globally, until the terrorist attacks on the United States which took place on 11th September 2001, when it drew attention to the importance of creating an independent OSINT agency to intensify exploiting these resources in order to protect national security. Some years later the U.S. government, merged the FBIS and other related research entities into one body. This organization is now called the Open Source Enterprise and it’s managed by the CIA. Nowadays, with the increase in volume of data available, government departments, non-governmental organizations (NGOs), and business corporations are starting to rely significantly on OSINT rather than private and classified information. OSINT sources differ from other forms of intelligence because they must be legally accessible by the public without breaching any copyright or privacy laws. For this reason, they are considered “publicly available.” This distinction makes the ability to gather OSINT sources applicable to more than just security services. For example, businesses can benefit from exploiting these resources to gain intelligence about their competitors.

OSINT includes all publicly accessible sources of information. This information can be found either online or offline, including in the following places:

- The Internet, which includes the following and more: forums, blogs, social networking sites, video-sharing sites such as YouTube.com, wikis, Whois records of registered domain names, metadata and digital files, dark web resources, geolocation data, IP addresses, people search engines, and anything else that can be found online.
- Traditional mass media (e.g., television, radio, newspapers, books, magazines)
- Specialized journals, academic publications, dissertations, conference proceedings, company profiles, annual reports, company news and employee profiles.
- Photos and videos including metadata
- Geospatial information (e.g., maps and commercial imagery products)

A vast amount of papers and books on the topic of OSINT can be found, as well as information regarding how and where to search for data about a specific target, thanks to OSINT research or specific OSINT tools. Two OSINT tools, Recon-ng and theHarvester, used during the development of the system that assesses reputation, are described below.


```

[recon-ng][default] > add domains
domain (TEXT): unicam.it
[recon-ng][default] > show domains

+-----+
| rowid | domain | module |
+-----+
| 1     | unicam.it | user_defined |
+-----+

[*] 1 rows returned
[recon-ng][default] > use recon/domains-hosts/
recon/domains-hosts/bing_domain_api          recon/domains-hosts/google_site_web
recon/domains-hosts/bing_domain_web         recon/domains-hosts/hackertarget
recon/domains-hosts/brute_hosts             recon/domains-hosts/mx_spf_ip
recon/domains-hosts/builtwith              recon/domains-hosts/netcraft
recon/domains-hosts/certificate_transparency recon/domains-hosts/shodan_hostname
recon/domains-hosts/findsubdomains         recon/domains-hosts/ssl_san
recon/domains-hosts/google_site_api        recon/domains-hosts/threatcrowd
[recon-ng][default] > use recon/domains-hosts/findsubdomains
[recon-ng][default][findsubdomains] > show info

Name: FindSubDomains DNS search
Path: modules/recon/domains-hosts/findsubdomains.py
Author: Pedro Rodrigues (@Pedro_SEC_R)

Description:
  Queries the FindSubDomain page for sub-domain information in a domain.

Options:
  Name      Current Value  Required  Description
  -----
SOURCE     default          yes       source of input (see 'show info' for details)

Source Options:
  default   SELECT DISTINCT domain FROM domains WHERE domain IS NOT NULL
  <string>  string representing a single input
  <path>    path to a file containing a list of inputs
  query <sql> database query returning one column of inputs

[recon-ng][default][findsubdomains] > █

```

Figure 4.2: Recon-ng module options

In this example the *unicam.it* domain was used along with the *findsubdomains* module to gather some subdomains. When 'run' is typed, Recon-ng will execute the module and the output of the research will be as shown in the following pictures 4.3.

The outcome is only a module that finds subdomains from the many that are present in Recon-ng. Each module has its own request in different servers to gather different information. Some modules also need an API key to retrieve data from certain servers such as Bing, Censys, GoogleCSE, IPInfoDB or Shodan.

```
[recon-ng][default][findsubdomains] > run
-----
UNICAM.IT
-----
[*] [host] root.camcic.unicam.it (<blank>)
[*] [host] elearning.unicam.it (<blank>)
[*] [host] rws3.unicam.it (<blank>)
[*] [host] missioni.unicam.it (<blank>)
[*] [host] dataentry.unicam.it (<blank>)
[*] [host] apperasmus.unicam.it (<blank>)
[*] [host] dida.cs.unicam.it (<blank>)
[*] [host] jobfinder.unicam.it (<blank>)
[*] [host] idp.unicam.it (<blank>)
[*] [host] myprint.unicam.it (<blank>)
[*] [host] docenti.unicam.it (<blank>)
[*] [host] idp.cs.unicam.it (<blank>)
[*] [host] vele.unicam.it (<blank>)
[*] [host] didattica.unicam.it (<blank>)
[*] [host] computerscience.unicam.it (<blank>)
[*] [host] conferences.cs.unicam.it (<blank>)
[*] [host] web.unicam.it (<blank>)
[*] [host] nutrition.unicam.it (<blank>)
[*] [host] rdeleone.sst.unicam.it (<blank>)
[*] [host] blog.unicam.it (<blank>)
[*] [host] cosy.cs.unicam.it (<blank>)
[*] [host] cuscamerino.unicam.it (<blank>)
[*] [host] gnxas.unicam.it (<blank>)
[*] [host] vela.unicam.it (<blank>)
[*] [host] musicultura.unicam.it (<blank>)
[*] [host] www.didattica.unicam.it (<blank>)
[*] [host] didanew.unicam.it (<blank>)
[*] [host] iceproject.cs.unicam.it (<blank>)
[*] [host] ulisse2.cs.unicam.it (<blank>)
[*] [host] serg.unicam.it (<blank>)
[*] [host] vetrouter.unicam.it (<blank>)
[*] [host] sunlb2.unicam.it (<blank>)
[*] [host] pcman3.unicam.it (<blank>)
-----
SUMMARY
-----
[*] 33 total (33 new) hosts found.
[recon-ng][default][findsubdomains] > █
```

Figure 4.3: Recon-ng “findsubdomains” module output

4.2.2 theHarvester

theHarvester [35] is a tool used for gathering subdomain names, email addresses, virtual hosts, open ports/banners, and employee names from different public sources such as Google, Bing, LinkedIn, Twitter, Yahoo, PGP keys, and more. The search conducted using this tool is passive, meaning that the target will not notice any reconnaissance activities from your side. This tool is intended to help Penetration testers in the early stages of the penetration test in order to understand the customer footprint on the Internet. It is also useful for anyone that wants to know what an attacker can see about their organization. theHarvester is a console tool and comes pre-installed on Kali Linux as well like Recon-ng, the picture 4.4 shows the list of commands that you can run. It’s possible to select different data sources such as Baidu, Bing, Google, GoogleCSE, LinkedIn, PGP, Twitter, vhost, VirusTotal, netcraft, Yahoo, and so forth. We can also perform active attacks, including DNS brute force attacks, DNS reverse lookups, and DNS Top-Level Domain (TLD) expansions.


```
[+] Emails found:
-----
rettore@unicam.it
protocollo@pec.unicam.it
roberta.tacchi@unicam.it
ortobotanico@unicam.it
elide.ferranti@unicam.it
tecnica@pec.unicam.it
pierluigi.palombi@unicam.it
dirittocivile@unicam.it
andrea.catorci@unicam.it
andrea.dallasta@unicam.it
ludovico.romagni@unicam.it
Ehiofomwana.omwanghe@studenti.unicam.it
synbiotec@unicam.it
ilo@unicam.it
medicina.estetica@unicam.it
scuola.dirittocivile@unicam.it
luca.tesei@unicam.it
orientamento@unicam.it
francesco.nobili@unicam.it
gianni.sagratini@unicam.it
lucia.ruggeri@unicam.it
sandra.pucciarelli@unicam.it
diego.campetella@unicam.it
emanuela.merelli@unicam.it
roberto.canullo@unicam.it
paolo.scrollavezza@unicam.it
andrea.morichetta@unicam.it
sfp@pec.unicam.it
cifani@unicam.it
carlo.bisci@unicam.it
piero.farabollini@unicam.it
silvia.vespasiani@unicam.it
```

Figure 4.5: theHarvester, email addresses found

```
[+] Hosts found in search engines:
-----

Total hosts: 23

[-] Resolving hostnames IPs...

.unicam.it:empty
2525252Fidp.cs.unicam.it:empty
2525252Fidp.cs.unicam.it:empty
253Dwww.unicam.it:empty
ANNO...unicam.it:empty
D7.unicam.it:94.177.192.171
Web.unicam.it:193.204.8.30
biblioteche.unicam.it:94.177.192.171
computerscience.unicam.it:94.177.192.171
d7.unicam.it:94.177.192.171
docenti.unicam.it:193.204.8.131
idp.cs.unicam.it:193.205.92.90
isas.unicam.it:94.177.192.171
juris.unicam.it:94.177.192.171
orientamento.unicam.it:94.177.192.171
pec.unicam.it:empty
polocivitanova.unicam.it:94.177.192.171
pros.unicam.it:193.205.92.119
studenti.unicam.it:empty
web.unicam.it:193.204.8.30
www.sst.unicam.it:94.177.192.171
www.unicam.it:94.177.192.171
www2.unicam.it:empty
```

Figure 4.6: theHarvester, hostnames found

4.3 Payable Services

All the information needed to assess the reputation cannot be found through the OSINT research, that's why some payable services were used during the development of the thesis. In the specific to complete the final report, the information regarding the financial, legal and social audit are needed. Specifically in order to complete the final report, information regarding the financial, legal and social audit are required. While carrying out the research online, two payable services named Cribis D&B and SocialNet were found which will be implemented in the system. Below, these two services and their uses are described.

4.3.1 Cribis D&B

Financial, legal and Business information about a company can be found on the website of "Camera di Commercio" [36]. This website contains a public register in which Italian companies, foreign companies with headquarters or local units in Italy and other entities (eg foundations, associations) that carry out an economic activity, are required to register.

The Business register is the primary source of certification of the constituent data of the registered economic subjects, as well as the municipal registers are for citizens' data. As written, this register is public on that website, so everyone can find information

about a specific company at any time. An huge amount of data was received from this website, so a service called Cribis D&B 4.7 was used to retrieve parsed information from that website.



Figure 4.7: Cribis logo

“Cribis D&B is a highly specialized company with advanced business information skills. It was set up with the aim of providing both the Italian market and D&B’s global clients with the highest standards of quality in terms of business coverage, the depth and accuracy of information, the technological capacity and flexibility to respond quickly to market demands, as well as in the provision of decisional systems, and scoring models” [37].

Cribis D&B guarantees the highest quality standards and the maximum coverage in economic and commercial information to companies all around the world. Cribis D&B collects millions of payment experiences in order to detect the payment habits of Italian and foreign companies towards them of their suppliers, considered individually or grouped by sector.

The Cribis service used provides a business report containing all the information needed to evaluate a company and get to know customers, prospects or suppliers in depth. Such information includes: credit scoring and commercial credit, all public data, annual budgets with the main performance indexes, information on payment behavior, corporate ties, investments ,membership, exponents, negative events, banks, management, economic activity and market and also local units.

The Cribis report service was implemented on this work thanks to their A2A technology using the XML protocol, this technology allows a report to be requested through the integration with their web services.

4.3.2 SocialNet

SocialNet [38] is another payable service that will be implemented to retrieve information about the social audit of our reputation index model. it is used to uncover Aliases & Relationships that Mirror the Physical World, it’s possible to uncover identities, correlations, networks of associates and available geographical information in just minutes. Since bad actors likely use the internet to communicate for themselves personally, to coordinate criminal activities or as a tool for malicious actions, SocialNet can be invaluable for both cyber or physical criminal investigations and social media forensics. SocialNet is a Maltego commercial transformation package (Maltego XL, Maltego Classic) that can be integrated into other platforms via Restful API.

4.4 Company and Third Party interaction

Europe, unlike the United States and the rest of the world, has strict laws and regulations on the management and manipulation of private data. This certainly leads to a more difficult if not impossible finding of some useful information for the creation of the final report. A solution can be found to refer to gaps in information that cannot be found during the data search phase.

Why not ask for information directly from the person concerned? Or in this case, from the defined target. Or why not ask for information from people or companies connected to the target, for example, its customers, suppliers or even employees?

The phase of requesting information from third parties must obviously be managed with the complete consent of the parties themselves. Despite requesting specific information, they will not always be willing to give it and, in some cases, perhaps the parties could also doubt what the intentions are with the information received. For this, an activity similar to a due diligence process will have to be carried out. A team will take care of the information request, analysing it and then inserting it into the system that will be created.

A very important aspect to take into consideration is the reliability factor of the information received from these third parties. Giving for consolidated that the information found online through paid tools and services are true, their degree of reliability will be equal to 100 percent, while for the information received from third parties their degree of reliability will obviously be lower. According with the Alfa Group Company it was decided to give a degree of reliability of 70 percent for the information received from third parties such as suppliers, employees, or collaborators, while for the information received from the target company itself the degree of reliability will be 50 percent.

Through this solution, it will be possible to receive all or part of the information that was not found online, and then create a final report containing the information for all or almost all of the components from the defined model.

Summary

The Internet provides an immense amount of information but, as already seen and described in this chapter, the possibility of finding personal information on the Web varies from state to state. Knowing that a possible target company will have to be Italian or at least European, a lot of information is protected by regulations and laws for privacy and therefore difficult or impossible to find. Through the use of some OSINT tools and payable services, however, some data is retrieved and used by the system that will be developed. While for all other information that is not possible to find, we will rely on the information given to us by third parties or by the target company itself.

With the different ways to gather information now being defined, the next chapter will describe how the final system was implemented, from the research and analysis phase to the creation of the report and the user interface.

5. Reputation System

While the previous chapters showed all the necessary data sources needed in order to build a system able to generate a reputation report, this chapter is going to explain the entire system built to achieve the goal of the thesis. From the index model reputation, to the developed method used to gather information, to the user interface, these are the main parts of the system described in this chapter.

While the previous chapters showed all the necessary data sources needed in order to build a system able to generate a reputation report, this chapter is going to explain the entire system built to achieve the goal of the thesis. From the index model reputation, to the developed method used to gather information, to the user interface, these are the main parts of the system described in this chapter.

In the project overview chapter the system requirements were discussed. All the tasks that the system should do were explained as well as, thanks to the information gathering chapter, where to find all the necessary information to create the report. Chapter 4 demonstrated that not all the information could be found, however the most important could through payable services and OSINT tools. Some methods were developed and described throughout this chapter.

Moreover, the user interface needed to interact with the created system will be described in detail, from the home page containing the form used to start a new research of a target, to all the sections useful to the user in order to interact with the system.

The system is essentially composed of two main parts: the client side and the server side. The client side containing the user interface and the server side containing the reputation index model index and the information gathering parts divided by OSINT tools and payable services. The next section describes the architecture of the system, showing how the client side and server side are connected, to then analyse and describe every part of the system throughout the chapter.

5.1 System Architecture

This section analyses the architecture of the system describing its organisation and also defining the interested parties connected to it. The idea is to create a system with a graphical interface where the user can enter some information about the target, namely the VAT number and website, and a server side that can handle all the functionalities defined in the chapter 3.1. In agreement with the Alfa Group Company it was decided to create a web based application whose server side worked in Java and the client side with classic HTML, all on a REST architecture. Having said that, the final choice ended up being JAX-RS.

REST is an architectural style that defines a set of constraints that, when applied to the architecture of a distributed system, induce desirable properties like loose coupling

and horizontal scalability. RESTful web services are the result of applying these constraints to services that utilize web standards such as URIs, HTTP, XML, and JSON. Such services become part of the fabric of the Web and can take advantage of years of web engineering to satisfy their clients' needs. The Java API for RESTful web services (JAX-RS) is a new API that aims to make development of RESTful web services in Java simple and intuitive [39].

JAX-RS is one of the latest generations of Java APIs that make use of Java annotations to reduce the need for standard base classes, implementing required interfaces, and out-of-band configuration files. Annotations are used to route client requests to matching Java class methods and declaratively map request data to the parameters of those methods. Annotations are also used to provide static metadata to create responses. JAX-RS also provides more traditional classes and interfaces for dynamic access to request data and for customizing responses [39]. The most popular implementations of JAX-RS used by developers are RESTEasy and Jersey.

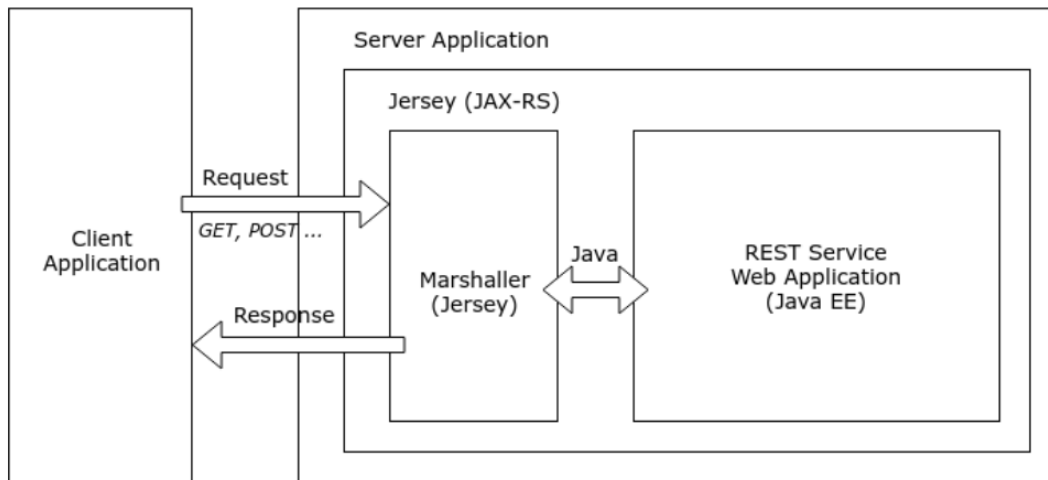


Figure 5.1: System Architecture

The picture 5.1 shows the system architecture composed by the client side containing the User Interface that will be described in the chapter 5.6, and the server side composed by JAX-RS with Jersey implementation running on Tomcat.

Apache Tomcat (or simply Tomcat) is an open source web server developed by the Apache Software Foundation. It implements JavaServer Pages (JSP) and servlet specifications, thus providing a software platform for running Web applications developed in Java language. Tomcat is a service that implements the Java EE specification such as servlets and JSPs.

Before developing the reputation system, some files such as *web.xml* and *pom.xml*, have been configured in order to let the entire architecture work. *Web.xml* specifies the JAX-RS servlet used to run the JAX-RS code and also provides configuration and deployment information for the web components that comprise a web application. The *pom.xml* file instead contains information about the project and configuration details used by Maven to build the project. It contains default values for most projects. When executing a task or goal, Maven looks for the POM in the current directory. It

reads the POM, gets the needed configuration information, and then executes the goal. The *pom.xml* file in this project contains all the dependencies about Maven, Jersey implementation, Apache Tomcat and JSON.

The whole architecture has been defined and explained. Now, how the system was built based on the architecture and the processes that it performs will be described. The client side is essentially composed by the user interface in which the user enters the data necessary to start the research. The server side analyses the request made by the client side and begins the data gathering phase, which will be described in the chapters 5.3 and 5.4. Then, the information parsing phase and recursive research will be carried out in order to generate a report. On completion of these actions, the system will execute the algorithm that calculates the reputational score of the target, described in the chapter 6, and return a response to the client side. If all the operations have been performed correctly, the user will be able to download and open the previously generated report.

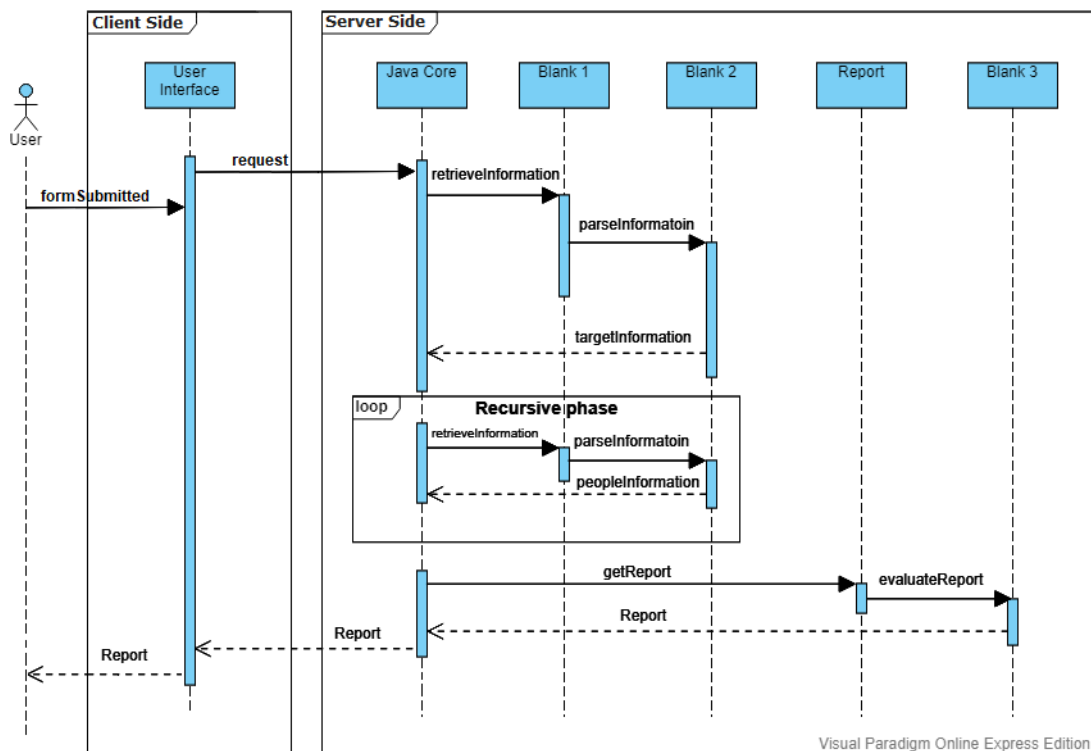


Figure 5.2: Process Diagram

The picture 5.2 shows the processes performed by the system, the components that manage the processes and in which order they are carried out. The client side only takes care of launching the search and waiting for a server result, while the server side takes care of most of the work. When the server receives a request from the client, it starts the information gathering phase on the chosen target, it parses the information found through the reputation index model and then starts the recursive phase on the people or companies connected to the target. It parses the information retrieved from the recursive phase and generates a report containing all the information processed. Once this is done, the server will carry out the evaluation phase of the report calculating

the reputational score of the target. For concluded operations the server will return a positive response to the client which will be able to be downloaded and the report created can be opened.

5.2 Reputation Index Model

Through the literature review, in the chapter 2, the corporate reputation and due diligence concepts were analysed. The corporate reputation literature has focused on a variety of definitions of the corporate reputation construct. This variety has resulted in a lack of agreement about what corporate reputation is, about its key components and about the relationship between them. Also, from the due diligence point of view, there are different ideas, from those considering it as just a process to find legal, financial and management information to those that expand this process to other concept such as ethical, social, marketing and products/services.

After examining and evaluating the literature on corporate reputation and due diligence, it emerged that a union between the two concepts was necessary to fully define all the components of the reputation according to this thesis, even if it became clear that some parameters were very similar, if not identical. Further indicators were then added on request of the Alfa Group Company, all concerning the technological components.

The merging between corporate reputation and due diligence parameters created the index model showed in the table 5.1. It illustrates the key components of the reputation index and includes parameters for each of the components, the key components are grouped into categories for easier use of the model.

Contrary to other models of reputation seen during the literature review, in which illustrative measures for each of the components were shown, in this model all the parameters that affect the definition of reputation are defined, both from internal and external information. If a parameter is not relevant, it is neither an element that improves the reputation nor a component that destroys the reputation and therefore will not affect the overall score, and therefore is not present in the list.

Table 5.1: The Reputation Index Model

Category	Component	Parameter
Business	General Information	Company Structure
		Overview Products / Services
		Overview Global Activities
		Corporate Vision and Mission
		CVs of Key Personnel
		Market Definition and Segmentation
		Memberships in Associations
		Intercompany Relationships
		List of Participations
		Strategy
	Competition Analysis	
	Market Research and Marketing Studies	
	List of Acquisitions and Disposals	
	Continued on next page	

Table 5.1 – continued from previous page

Category	Component	Parameter
Business	Sales and Marketing	Description Sales Organization
		Description Marketing Organization
		Sales Literature and Forms
		Standard Customer Contracts
		List of 20 largest Customers/Channels
		Sales Alliances
		Backlog Development
		Company Publications
		Customer Complaint Reports and Management
		Customer Satisfaction Surveys
		Operations
	Procurement Organization and Process	
	Major suppliers	
	Form of Standard Purchase Order	
	Quality Management	
	EHS Management System	
	Desaster Recovery Plans	
	R&D	Description R&D Organization
		R&D Strategy and Roadmap
	IT	Overview IT Organization
		List of pending IT Projects
		Overview of Software used
	Environmental	Environmental Reports
		List of Hazardous Materials handled
Schedule of Incidents		
Ethical	Ethical balance sheet	
Legal	Corporation	Legal Company Structure
		Competition
		Corporate History
		Lists of all current Shareholders
		Cap Table
		Samples of Common and Preferred Stock Certificates
		Stock Option Plans
		Convertible Debt Agreements
		Copies of any Voting Trust Agreements
		Trade Register Excerpts
		Business Registration
		Licences, Permits and Certificats
		Shareholders' Agreement
		Articles of Incorporation
		By-laws
		Partnership or JV Agreements
		Securities Issuances
List of all Officers and Directors		
Reporting to the BOD and Shareholders		
Minutes of Meetings		

Continued on next page

Table 5.1 – continued from previous page

Category	Component	Parameter	
Legal	Contracts	Material Supplier Agreements	
		Material Customer Agreements	
		Licensing and Reseller Agreements	
		List of outstanding Leases	
		Guarantees given by the Company	
		Credit Agreements	
	Disputes	Current Litigations	
		List of legal disputes	
		List of tax disputes	
	Compliance	Description of Risk Management System	
		Restrictions of doing Business	
		Compliance with Licences, Permits and Certificates	
		Business with Embargo Countries	
		Notices received from government organizations	
	Intellectual Property	Patents, Trademarks, Copyrights, Domaines	
		Out-licensing Technology Agreements	
		In-licensing Technology Agreements	
		Software Escrow Agreements	
		Sponsored Research Agreements	
	Insurance	List and details of Insurance Policies	
		List of past material Insurance Claims	
	Real Estate	Deeds	
		Leases of Real Property	
Copies of Appraisals			
Black lists	Presence in Financial Black list		
	Presence in Technologies Black lists		
Financials	Financial System	Overview Accounting System	
		Accounting Principles	
		Overview of Reporting Structures, Departments	
		Annual Planning and Forecasting Procedures	
		Treasury	
	Financial Statements	Annual Financial Statements - consolidated	
		Annual Financial Statements - Companies	
		Managements Correspondence with Auditors	
		Top 20 Accounts Payable	
		Top 20 Accounts Receivable	
		List of fixed Assets	
		List of financial Investments and Securities	
		Bad Debt Development	
		Details on Year-End Provision and Accruals	
	Management Reporting	Current Management Reporting	
		Interim financial Statements	
		Current Budget and Forecast	
		Historic Budget vs. Actuals	
	Business Plan	Current Business Plan	
		Details planned Capital Expenditures	
			Continued on next page

Table 5.1 – continued from previous page

Category	Component	Parameter
Financials	Financing	Bank Accounts and short-term Investments
		Loan Agreements and other Debt Instruments
		Finance and Operating Lease Agreements
	Taxation	Tax Accounts
		Tax Returns
		Notifications by Tax Authorities
HR	Overview	List and Details on Employees
		List of HR Litigations
		Invention Assignment Agreements
		Non Compete Agreements
	Compensation	List and Details of Compensation
		Bonus Plans, Retirement Plans, Pension Plans
		Owners Compensation
		Valuation Report Pensions
	Contracts	Contracts of Key Employees
		Standard Working Contracts
Digital	Website	Website Analysis
		Servers Location
	E-mails	Staff E-mails
		Check of penetration (breach)
Social	Stakeholder	Evaluation of stakeholders on products/services
		Feedback
	Social Network	Facebook analysis
		Linked-In analysis
		Twitter analysis
		Instagram analysis

As it's possible to see from the table 5.1, there are six main categories into which the model has been divided. The first one, "Business", includes general information about the company as well as the strategy of the company which analyses the strategic plans and the competition with other companies. Moreover, this category includes the sales and marketing fields assessing the company sales and marketing strengths and weaknesses, in addition to the operations of the company. The remaining and smaller sub-categories of the "Business" category include R&D, IT, environmental and ethical and they evaluate research and development organization and strategies, the companies IT projects and software used, environmental hazards and reporting and the ethical balance sheet, respectively.

The "Legal" category, instead, assesses the legalities, agreements and disputes of the company. This category is broken down into the following seven sub-categories: corporation, contracts, disputes, compliance, intellectual property, insurance and real estate. The corporation section considers corporate history, shareholders, competition, licenses and some legalities of the company. The contracts, disputes and compliance sub-categories include information regarding agreements with buyers and sellers, legal or tax disputes and conformities of the company, respectively. The final three sub-categories assess the intellectual property of the company, insurance policies and claims as well as appraisals and information regarding real property of the company.

The “Financials” category includes the financial system of the company, assessing a general overview of how the financial aspects of the company are managed. Following on from the financial system sub-category is the financial statements section which analyses annual statements, investments and accounts connected to the company. Management reporting, instead, assesses the current financial situation and compares it to the past situation. Moreover, sub-categories like business plan, financing and taxation analyse future plans for the company, financial agreements and information regarding the company’s tax.

The “HR” category includes a general overview of the HR department, what it deals with and how it works. It also contains information regarding the compensation of the company, evaluating employee rewards, pensions and employee contracts.

The “Digital” category, instead, analyses information regarding the website and e-mails throughout the company and checks also possible data breaches.

Finally, the “Social” category evaluates stakeholders on products and services, and examines the social networking of the company.

This model thus defines all the parameters involved in the definition of reputation. The illustrative measures for each key component will be defined and described in the 6 chapter, in which an algorithm to give a score on the target will be created.

5.3 OSINT tools modules Implementation

This section provides an overall description of how the OSINT tools shown in the 4.2.1 and 4.2.2 sections were implemented in the system, from the OSINT tools module analysis, to the Java Classes definition, to the requests made on the data sources.

As seen in the 4.2.1 and 4.2.2 sections, these tools provide a lot of different categories of modules. Their usage allows us to retrieve information such as emails, subdomains, contacts, credentials, social media profiles, and other information such as IP, reverse IP, WHOIS information, and ports information through different kind of modules. An analysis of which kind of modules was necessary in order to understand which were useful for our system. Based on the key components model index created in the previous section 5.2, the entire “Digital” component with its parameters can be researched.

Information regarding “Web analysis” can be found using the vulnerabilities modules of Recon-ng such as: *“punkspider.py, xssed.py and xssposed.py”*, these modules are useful in checking if a target website could be exposed to cyber-attacks. They make a request to different data sources checking if the website is present on the data source.

“Servers location” information can be retrieved in different ways. First of all sub-domains of the website should be searched, then using a Recon-ng module, the research of where the servers are can be carried out. Both tools were used in the sub-domains discovering phase, through the Recon-ng. The modules analysed were *“findSubDomainsHackertarget.py and findSubDomains.py”*. Instead, on theHarvester, the requests made in different Search Engines, such as Google, Bing, Yahoo and Baidu, were analysed. The information retrieved was then parsed, and a list of sub-domains could be returned. Now, using the *“findLocation.py”* module of Recon-ng, the information such as: IP address, country, region, longitude and latitude can be discovered.

Knowing that some emails can be found on the target website, more detailed research is carried out in order to retrieve as many emails as possible before checking if they are exposed to possible penetration through Data Breach. The term “Data Breach” refers to the intentional or unintentional release of secure or private/confidential information

to an untrusted environment. The phase of retrieving emails associated to the target is carried out with theHarvester in which, like in the sub-domains discovering phase, some requests are made in different Search Engines. Meanwhile data breaches are checked through the *“haveibeenpwned.py”* module of Recon-ng.

From the system programming point of view, a reverse engineering of the code of the modules described before was done. Since the OSINT tools are written in Python, a reverse coding from Python to Java has been done. Afterwards, an analysis of which modules are needed, a full understanding of how they work from a coding point of view and where they make requests to different sources are necessary to know. The code shown below 5.1, is an example of how a Python module is written on these OSINT tools. Those underlined show the google module, present on theHarvester, which retrieves all the information about a target such as emails, hostname, profiles and files through the Google Search Engine.

```

1 from theHarvester.discovery.constants import *
2 from theHarvester.parsers import myparser
3 import requests
4 import time
5
6 class search_google:
7
8     def __init__(self, word, limit, start):
9         self.word = word
10        self.results = ""
11        self.totalresults = ""
12        self.server = 'www.google.com'
13        self.dorks = []
14        self.links = []
15        self.database = 'https://www.google.com/search?q='
16        self.quantity = '100'
17        self.limit = limit
18        self.counter = start
19
20    def do_search(self):    #Scraping function
21        try:
22            urly = 'http://' + self.server + '/search?num=' + self.quantity + '&
23            start=' + str(self.counter) + '&hl=en&meta=&q=%40\'' + self.word + '\''
24            except Exception as e:
25                print(e)
26            try:
27                headers = {'User-Agent': googleUA}
28                r = requests.get(urly, headers=headers)
29            except Exception as e:
30                print(e)
31            self.results = r.text
32            if search(self.results):
33                time.sleep(getDelay() * 5)
34            else:
35                time.sleep(getDelay())
36            self.totalresults += self.results
37
38    def get_emails(self):    #Get emails function
39        rawres = myparser.Parser(self.totalresults, self.word)
40        return rawres.emails()
41
42    def get_hostnames(self):    #Get hostnames function
43        rawres = myparser.Parser(self.totalresults, self.word)
44        return rawres.hostnames()
45
46    def get_profiles(self):    #Get profiles function
47        rawres = myparser.Parser(self.totalresults, self.word)
48        return rawres.profiles()

```

```

49 def process(self, google_dorking): #Main function of this class
50     if google_dorking is False:
51         while self.counter <= self.limit and self.counter <= 1000:
52             self.do_search()
53             print(f'\tSearching {self.counter} results.')
54             self.counter += 100
55     else: # Google dorking is true.
56         self.counter = 0
57         print('\n')
58         print('[-] Searching with Google Dorks: ')
59         self.google_dork()

```

Listing 5.1: theHarvester Python Google module

A brief description of the code above 5.1 is necessary. Everything is managed by a Python class which first launches the *process* function, which then recalls *do_search* in a kind of loop. The *do_search* function retrieves all the information from the target domain through the parameter *urly*. A parse of the information is then carried out by the other functions defined in this class, such as *get_emails*, *get_hostnames* and *get_profiles*.

The useful part of the code needed to implement this script in java is, as well as the request made on the search engine, how theHarvester parses the information. The code 5.2 shows, in fact, different patterns applied on the retrieved information. The functions will return the effective information needed to be used in the report.

```

1 import re
2
3 class Parser:
4
5     def __init__(self, results, word):
6         self.results = results
7         self.word = word
8         self.temp = []
9
10    #Function that apply a pattern to the text to find emails
11    def emails(self):
12        self.genericClean()
13        reg_emails = re.compile(
14            r'[a-zA-Z0-9.\-+#~!$&\'",;=:]+ + '@' +
15            '[a-zA-Z0-9.-]*' + self.word)
16        self.temp = reg_emails.findall(self.results)
17        emails = self.unique()
18        return emails
19
20    #Function that apply a pattern to the text to find urls containing a file
21    def fileurls(self, file):
22        urls = []
23        reg_urls = re.compile('<a href="(.*?)"')
24        self.temp = reg_urls.findall(self.results)
25        allurls = self.unique()
26        for x in allurls:
27            if x.count('google.com') or x.count('search?hl'):
28                pass
29            else:
30                urls.append(x)
31        return urls
32
33    #Function that apply a pattern to the text to find hostnames
34    def hostnames(self):
35        self.genericClean()
36        reg_hosts = re.compile(r'[a-zA-Z0-9.-]*\.' + self.word)
37        self.temp = reg_hosts.findall(self.results)
38        hostnames = self.unique()
39        return hostnames

```

```

40
41 #Function that apply a pattern to the text to find people on LinkedIn
42 def people_linkedin(self):
43     reg_people = re.compile(r">[a-zA-Z0-9._-]* \\\ LinkedIn')
44     self.temp = reg_people.findall(self.results)
45     result = []
46     for x in self.temp:
47         y = x.replace(' | LinkedIn', '')
48         y = y.replace(' profiles ', '')
49         y = y.replace('LinkedIn', '')
50         y = y.replace('\"', '')
51         y = y.replace('>', '')
52         if y != " ":
53             result.append(y)
54     return result

```

Listing 5.2: theHarvester Python Parser Class

There are many different classes and modules analyses, the ones above are just some examples. All the modules have basically the same structure. The function that makes a request on a specific server and the functions that parse the retrieved information are analysed in each module.

Several Java classes have been developed, from the main one that handles all the requests to the one that parses the data found, obviously passing through the Recon-ng and theHarvester classes that contain the respective modules. The code 5.3 shows the different requests made in different data sources.

```

1 public class TheHarvester {
2     private static List<String> arrEmails = new ArrayList<String>();
3     private static List<String> arrSubDomain = new ArrayList<String>();
4     private static List<String> arrPeople = new ArrayList<String>();
5     private static String domain;
6
7     public TheHarvester(String domain) {
8         this.domain = domain;
9     }
10
11     /* Function that starts the research on different data sources */
12     public static void run() throws Exception {
13         baiduSearch(domain); /* Baidu Search Engine */
14         bingSearch(domain); /* Bing Search Engine */
15         googleSearch(domain); /* Google Search Engine */
16         yahooSearch(domain); /* Yahoo Search Engine */
17         linkedinSearch(domain); /* LinkedIn Social Network*/
18     }
19
20     /* Request on Google Search Engine */
21     public static void googleSearch(String domain) throws Exception {
22         String response = null;
23         String userAgent = "Mozilla/5.0 (Windows; U; Windows NT 6.0;en-EN; rv:1.9.2)
24             Gecko/20100115 Firefox/3.6";
25         Map<String, String> map = new HashMap<String, String>();
26         map.put("User-Agent", userAgent);
27         map.put("Host", "www.google.com");
28         int count = 0;
29
30         while (count < 1000) {
31             URL url = new URL("http://www.google.com/search?num=100&start="+count+"&hl
32                 =en&meta=&q=%40%22"+domain+"%22");
33             response = response + Utility.GETRequest(url, map);
34             count = count + 100;
35         }
36         arrEmails.addAll(Utility.findEmails(response, domain));
37         arrSubDomain.addAll(Utility.findSubDomains(response, domain));
38     }
39 }

```

```

37  /* Request on LinkedIn Social Network */
38  public static void linkedinSearch(String domain) throws Exception {
39      String response = null;
40      String userAgent = "Mozilla/5.0 (Windows; U; Windows NT 6.0;en-US; rv:1.9.2)
41      Gecko/20100115 Firefox/3.6";
42      Map<String, String> map = new HashMap<String, String>();
43      map.put("User-Agent", userAgent);
44      map.put("Host", "www.google.com");
45      map.put("Cache-Control", "no-cache");
46
47      int count = 0;
48      URL url = new URL("http://www.google.com/search?num=100&start=0&hl=en&meta=&
49      q=site%3Alinkedin.com/in%20unicam.it");
50      response = response + Utility.GETRequest(url, map);
51      arrPeople.addAll(Utility.findPeopleLinkedIn(response, domain));
52  }

```

Listing 5.3: theHarvester Java Class

What has been developed, reproduces the same behaviour as the OSINT tools modules and therefore the same results. In each function there is a HTTP request to a server in order to retrieve data, then the data is parsed through pattern and put in the global variable defined on the top of the code 5.3. An example of the results from the requests are shown in the pictures of the Appendix A.

All the information found during this first phase is inserted into the final report under the “Digital” key component. The next chapter is going to show how information, regarding all the other key components of the model defined, is being founded.

5.4 Payable service interactions

This section describes how the information has been retrieved through the Cribis service cited in the 4.3.1 section. Once the Cribis web services are portrayed, java classes created are analysed to fully understand how this service works.

Cribis is a payable service used in this system to gather information such as credit scoring and commercial credit, annual budgets with the main performance indexes, information on payment behaviour, corporate ties, investments, membership, exponents, negative events, banks, management, economic activity and market and also local units. The retrieving information phase was carried out using their Web Services, described through their WSDL.

According to [40], WSDL is an XML-based interface description language that is used for describing the functionality offered by a web service. The acronym is also used for any specific WSDL description of a web service, which provides a machine-readable description of how the service can be called, what parameters it expects, and what data structures it returns. WSDL is often used in combination with SOAP and an XML Schema (XSD) to provide Web services over the Internet. A client program connecting to a Web service can read the WSDL file to determine what operations are available on the server. Any special datatypes used are embedded in the WSDL file in the form of XML Schema. The client can then use SOAP to actually call one of the operations listed in the WSDL file using for example XML over HTTP.

According to [41], XSD (XML Schema Definition) is a World Wide Web Consortium (W3C) recommendation that specifies how to formally describe the elements in an Extensible Mark-up Language (XML) document. This description can be used to verify that each item of content in a document adheres to the description of the element in which the content is to be placed. XSD can also be used for generating XML

documents that can be treated as programming objects. In addition, a variety of XML processing tools can also generate human readable documentation, which makes it easier to understand complex XML documents.

Cribis WSDL contains explicitly the XSD with the definition of the XML schema both of the calls and of the answers exposed by the WEB Service. Cribis Web service allows the purchase and reconsultation of reports and the request for further information. The interaction to retrieve a report is a Request / Response type like shown in the picture 5.3.

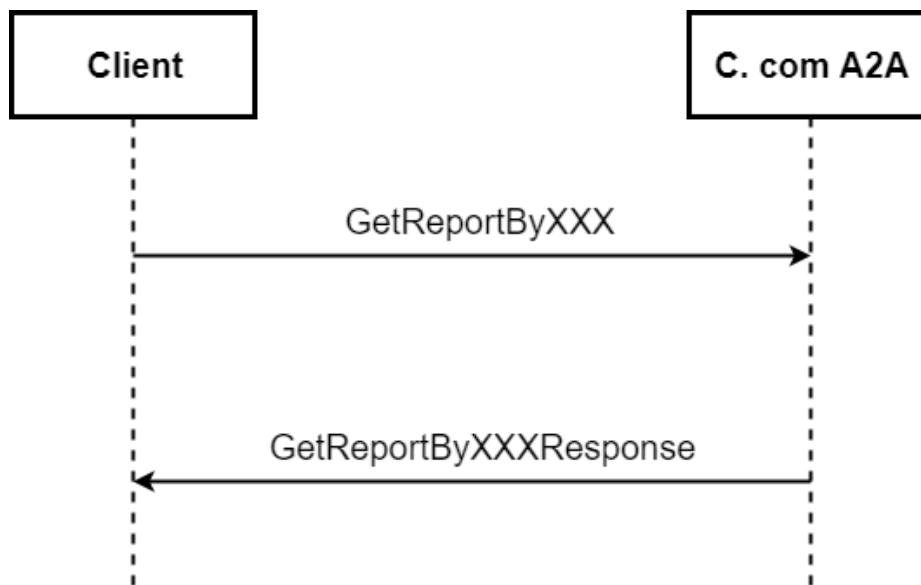


Figure 5.3: Cribis Request / Response model

Cribis allows requests to be made to two different main parties, companies and people. To purchase a business company report the following methods can be used:

- **GetReportByCCIAARegistration**, to retrieve a report specifying the REA number and the province of the CCIAA of the company sought.
- **GetReportByTAXorVATCode**, to retrieve a report specifying the tax code or the VAT number of the company sought.
- **GetReportByTAXCode**, to retrieve a report specifying the tax code of the company sought.
- **GetReportByVATCode**, to retrieve a report specifying the VAT number of the company sought.
- **GetReportByCRIFNumber**, to retrieve a report specifying the Cribis Number (unique identifier of a company assigned by Cribis) of the company sought.

All these methods have an identical input apart from the company identification parameters (TAX Code, VAT Code ... etc) specified on the request such as Username

and Password of the Cribis account, Transactio ID, Product ID and the product format output (XML or PDF).

Instead to purchase a person report the following methods can be used:

- **GetReportByPersonData**, to retrieve a report specifying Name, Surname, Date of birth, Place of birth, Province birth code of the person sought.
- **GetReportByPFTaxCode**, to retrieve a report specifying Name, Surname, Tax ID of the person sought.

The person report provides the outcome of the checks on the presence of negativity (protests, prejudicial and insolvency proceedings) on the single person questioned, both for individuals and legal entities.

In our system the two methods implemented are *GetReportByTAXorVATCode* for companies and *GetReportByPFTaxCode* for people. Through these requests it is possible to retrieve an XML regarding our system or a Java object containing all the information. The Java objects are structured like the XML thanks to the XSD. An analysis of the Company report schema with its elements is shown in the picture 5.4.

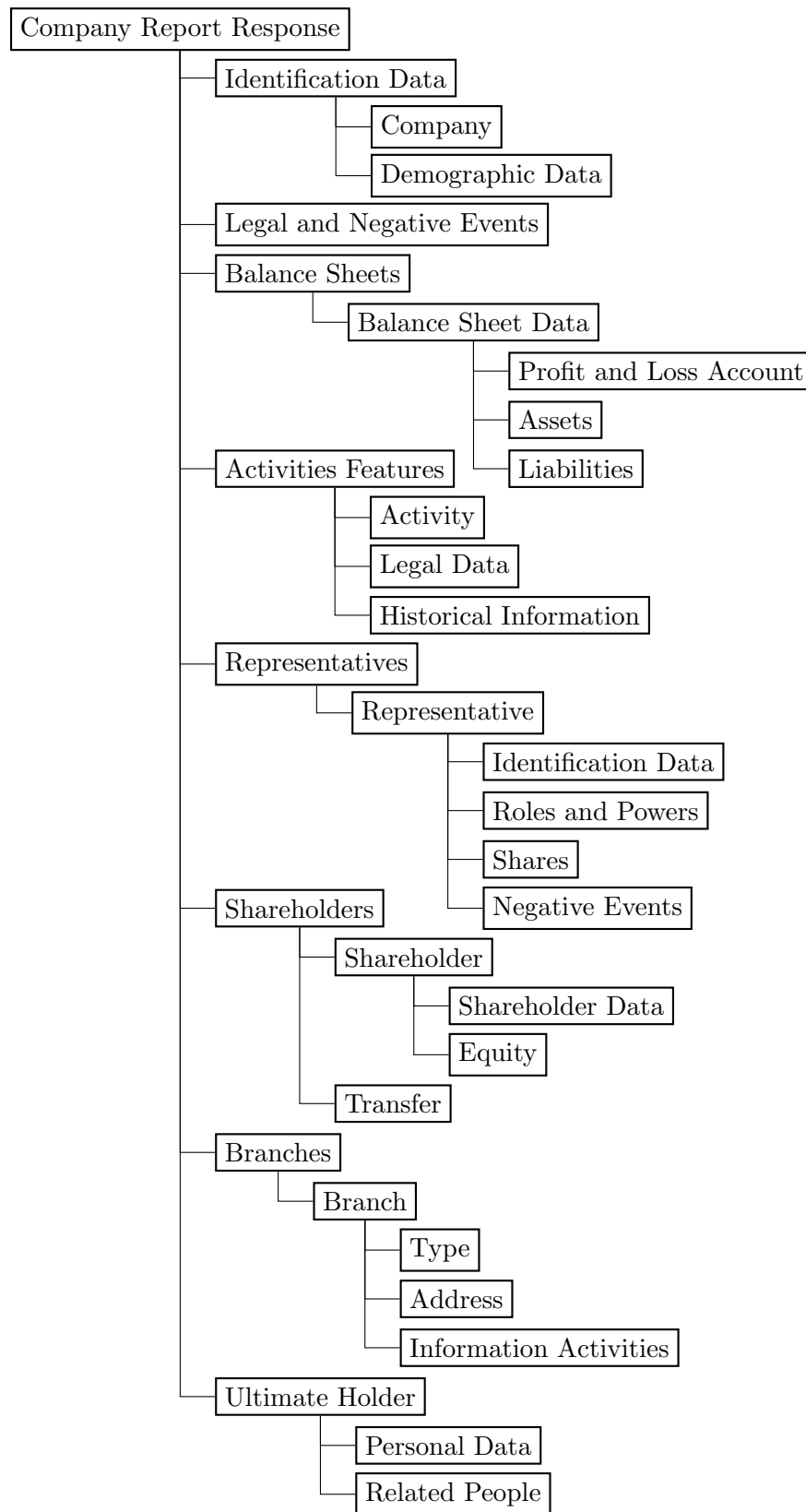


Figure 5.4: Cribis XML Schema example

It describes only the most important elements of the business company report XSD with which it was possible to retrieve only the information necessary for the report listed in the model. Of course a different XSD is built for a person report. From a programming point of view, the WSDL are converted into Java classes and the methods described before were implemented. The code 5.4 shows the request made by the system in order to retrieve a Java object containing all the information found on Cribis.

```
1 import _2012_04_20.holders.*;
2 import _2012_04_20.*;
3 import _2012_04_01.crif_cribiscom_Report_Company.holders.*;
4 import _2012_04_01.crif_cribiscom_Report_Company.*;
5 import _2012_04_01.*;
6
7 public class Cribis_Requests{
8
9     public static void companyMethod(){
10
11         BusinessCompanyReportProxy companyReport = new BusinessCompanyReportProxy
12         ();
13
14         companyReport.getReportByTAXorVATCode(username , password ,
15         applicationTransactionID , productID , productFormats , customerReferenceText ,
16         TAXorVATCode , customerInternalData , AAInternalData , checkPA ,
17         transactionResponse , businessReportDocument , businessReportAttachments ,
18         inquiryDetails , productDetails);
19     }
20 }
```

Listing 5.4: Cribis Business Company Report Method

The *getReportByTAXorVATCode* has a lot of parameters, some of them were explained before and all the others instead are optional. Once the code above is run, different queries on the structure elements are carried out to retrieve only the required information.

5.5 Recursive phase and Report generation

Using the OSINT tools methods and the payable service described before, it was possible to retrieve as much information as needed to assess the reputation of the target. The Java core system manages the gathering information phase as well as the recursive phase and the report creation phase.

Once the gathering information phase has been completed, a list of contacts, employees, managers or other people who this company are dealing with will be retrieved. Using this list, the system will perform a recursive research that will return a list of people or companies for each of the elements of the first list. The recursive research is performed with the same tools and services described before.

The recursive phase can be figured out such as a tree graph, in which at the top there is the main target (the target company) and below it, all the people or companies connected. The picture 5.5 below explains this concept better.

The recursive phase can be structured in layers in which the first one contains the people and companies directly connected to the target. The second one contains the people and companies connected to the layer above and so on. To make a deeper recursive research means having more layers in the tree graph and therefore have a wider list of people connected to the target. According to the Alfa Group Company,

the recursive phase will be carried out by the system three times. This means that the tree graph has a depth of three. Having a tree graph too deep would be useless because too many people or companies would be found.

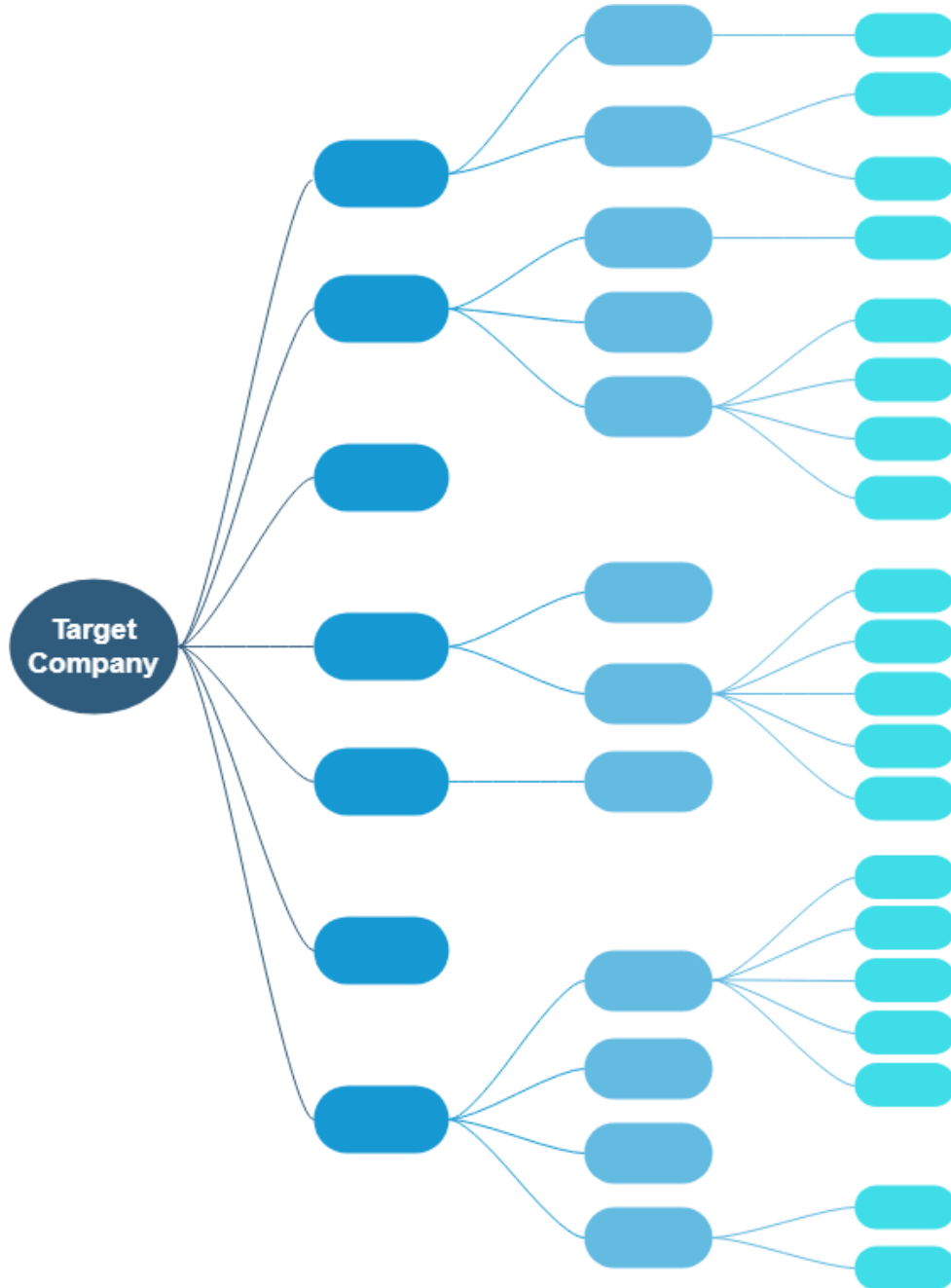


Figure 5.5: Network of people related to the Target Company

Considering that all people or companies connected to the target could have an agreement with it, the reputation of the target company can change also based on the reputation of the people or companies connected to it. For example, knowing that an employee has a legal liability, the reputation of the target company will be different than before.

The recursive gathering information phase is done in order to have a better idea

of all the possible people the target company is dealing with or it was dealing with, modifying, definitely, the target company reputation. The way in which the system will assess the reputation of all the people connected to the target will be the same one of the target company, but with less parameters to check.

Concluding the recursive phase, all the required information has been retrieved, so a report for the user can be created. Following the structure of the model created in the previous chapters, the system will build a long report containing, first all the information about the target company sorted by key components, then a list of connected people with their score of reputation as well.

The report can be downloaded from the user interface of the system once the gathering information phase is finished. An example of the report is shown in the list of pictures in the Appendix C.

5.6 User Interface

After having explained the main functionalities of the system, the last but not least important component is the User Interface (UI). It is responsible for launching a search on a target company and retrieving the report generated as a PDF. Through a user-friendly interface, it is possible to surf in different sections, from the one needed to launch a research phase on a target, to the ones in which it is possible to insert information retrieved by the same target company or by a third party, to the one that shows the network of people connected to the target.

A more in depth description of all these sections is necessary to better understand how the user interface works. As noticed in the system architecture chapter, the system has a web-based interface that communicates with Java classes in order to launch different functionalities of the system. All the pictures regarding the user interface are shown in the Appendix B.

The UI home page is shown in the picture B.1, in which there is a form used to start the gathering information phase. As mentioned in the previous chapter, the information required to conduct a research is the company name, the VAX or VAT code and also, if present, the company website. Once the “Search” button is submitted, a request to the main java classes is made and a gathering information phase will start. The UI will be on the loading screen during the entire phase.

As described in the section 4.4, most of the necessary information to assess the reputation cannot be found only through the OSINT tool or payable services, therefore, there are two UI sections regarding the possibility to manually add information received from the same target company or a third party such as suppliers, employees, or collaborators. The picture B.2 shows the company section, in which there is a form which requires information to be inserted, sorted by categories the same as in the reputation index model. In the Business tab, for example, the information requested by the systems are the presence of R&D Strategy plan in the company, the ethical balance sheet, the joint venture agreements and the presence of the liability claims.

The UI provides a network section in which a kind of tree graph of people connected to the target company is shown. In fact the concept is the same as in the picture 5.5, in which different layers of people or companies connected to the target are shown. The picture B.3 shows the network section of the user interface where there is an interactive network of nodes. Every node shows the VAT or VAX code of the company or person, while the edge shows the role of that node based on the previous node.

The last section of the user interface provides the opportunity to download and open the generated report on the defined target.

Summary

The entire system, and how it works, has been described, from the user interface that allows the use of the system, to the Java program that runs on the server. The initial literature review gave the foundations to the thesis to design a model index containing all the components necessary to assess the reputation. Through this model, the system will parse the information retrieved and also it will create the report using the same structure.

Once a target company is submitted into the user interface home page, the main Java class starts to make requests in different data sources in order to retrieve as much information as possible. The data retrieved will then be parsed based on the index model and then stored in a report to keep the information available.

During the initial literature the key components of the reputation and how to measure the reputation using different weights for each key component were shown. In the next chapter, how the system is going to measure the reputation using the defined model is going to be described. The different weights given to the key components will be described based on the different stakeholders described in the chapter 3.2.

An algorithm that will calculate a reputational score of the target will be described and an example for each stakeholder involved in the thesis will be shown. Moreover a reliability score of the reputational score will be calculated.

6. Reputation Evaluation Algorithm

As already mentioned, the main objective of this thesis is to design and implement a system that creates a report containing the information necessary to assess the reputation. As well as this, the report will have to be evaluated by an algorithm in order to give a reputational score of the defined target. The reputation score will be used to understand the degree of reliability of the target; the higher the reputational score, the more reliable the chosen target will be.

For this reason, the purpose of this chapter is to explain how the algorithm that evaluates the information contained in the report was created, starting from which paper inspired it, then describing the necessary measures to evaluate the categories of the model defined in chapter 5.2, then moving to the weights given to the different categories based on the stakeholders shown in chapter 3.2. Finally, the implementation of the algorithm on the system will be shown.

The initial review of the literature was useful in the development of this thesis, providing not only the definition of the corporate reputation and due diligence with their key components, but also showing how it's possible to assess the reputation through the key components. The paper [30] shows, in fact, a list of measures for each key component, based on an anchor scale from 1, as the lowest or least desirable value, to 10, as an ideal value. After the measures have been averaged for each separate component, a weight must be given for each component, so that a weighted average can be calculated to create a score showing the degree of reputation.

Considering this paper [30] as an excellent model from which to take inspiration, a list of measures will be defined in order to create a reputation evaluation algorithm. The model created in the chapter 5.2, defines the key components and their parameters that are used to define the reputation. Instead, the measures that will be defined in this chapter will be used to evaluate the degree of reputation of the target.

6.1 Measures of Reputation Index Model

The created model defines all the key components and parameters required to evaluate the concept of reputation, but it does not provide the necessary measures for the algorithm to give a reputation score to the chosen target. For this reason, measurements have been defined with a relative scale of values. Since there is no possibility to give measures for all the key components and parameters of the model, a list of measures to assess the degree of reputation has been defined on the basis of each category of the model itself.

Table 6.1 illustrates the categories of the reputation index model and includes measures for each of the categories. Similar to in the paper, an anchor scale is given for each measure in order to transform them into a form more suited to computing a reputation score. The scale values for the measures in the table are anchored with 10 as an ideal for the company and 1 as the lowest, or least desirable for the company.

According to the Alfa Group Company, the following measures are defined for each category. Having created a completely automatic system, it was not possible to define measures that would analyse the information contained in the report, for this reason, as shown in the table 6.1, most of the measurements refer to the presence of a specific object or entity. In the case of a presence, a count of the quantity can be made and therefore a value can be assigned to the measurement.

The Business category checks the presence of strategic plans, management systems, internal R&D department and environmental rules. Moreover, it checks the presence of eventual accidents or ethical policies. The Legal category checks, instead, the presence of certificates, authorizations and licenses, legal or tax disputes, activities with countries subject to embargo, insurance policies and the presence in black lists. Moreover the feedback of the main partners will be evaluated, as well as the communications with government agencies. The Financial category assess the financial statements through the most important indicators such as ROI, ROE and ROS. The return on investment (ROI) is a balance sheet index that indicates the profitability and economic efficiency of the typical management regardless of the sources used: it expresses, that is, what makes the capital invested in that company. The return on equity (ROE) is a measure of the profitability of a business in relation to the equity, also known as net assets or assets minus liabilities. It is a measure of how well a company uses investments to generate earnings growth. The Return on Sales (ROS) makes it possible to assess the company's profitability in relation to the turnover achieved, by relating operating profitability and revenues. The Digital category, instead, checks the presence of breach in their system and of possible vulnerabilities on their website. Moreover an evaluation of the website server location will be carried out. The Human Resources category checks the presence of nominations for open positions in the company and the disputes or internal problems between employees. Finally, the Social category assesses the feedback from the customers.

Notice that the lack of some of these objects or entities in the report can be very relevant for the purpose of calculating the reputation score. Therefore, their non-presence is already in itself a value, which applied to some measures can be a critical element able to condition the entire reputation score. For example, not finding the balance sheets in the report, the reputation level will be with the minimum value because the target cannot be a candidate for any position without the relative balance sheets.

The responses to the scales for the individual measures should be averaged for each separate category. Once there is an aggregate measure (value of 1–10) for each category, an overall measure can be created by applying weights to each of the categories and summing the values. The next section is going to show the different weights applied per candidate described in the chapter 3.2.

Category	Measures	Anchor scale values:	
		least desirable	optimal (ideal)
Business	Presence of strategic plans	None	Numerous
	Product/Service quality associations	None	Numerous
	Presence of management systems	None	Numerous
	Presence of internal R&D	None	Numerous
	Presence of environmental rules	None	Numerous
	Claims for compensation	Several	Never
	Ethical policies	None	Numerous
	List of accidents	Several	Never
Legal	Market stability	Unstable	Long-term stability
	Presence of certificates, authorizations and licenses	None	Numerous
	Feedback from the main partners	None	Highest
	Presence of JV	None	Several
	Presence of legal or tax disputes	Numerous	Never
	Presence of activities with countries subject to embargo	Several	Never
	Evaluation of communications with government agencies	Negative	Positive
	Presence of insurance policies	None	Numerous
	Presence of patents, trademarks, copyright	None	Numerous
	Presence in black lists	Several	Never
Financial	Assessment of financial statements through ROE indicators	Low	High
	Assessment of financial statements through ROI indicators	Low	High
	Assessment of financial statements through ROS indicators	Low	High
Digital	Website server location	Uncommon places	Common places
	Presence of breach	Never	None
	Presence of vulnerabilities on the Website	Numerous	None
HR	Employee satisfaction	Quite low	Highest
	Nominations for open positions	None	Numerous
	Disputes or internal problems between employees	Numerous	None
Social	Feedback from the customers	Negative	Positive

Table 6.1: Measures of the Reputation Index Model

6.2 Assigning Values to the Index Model Categories

Putting stakeholders together on a scale of fair weights is certainly the hardest thing to do. Considering the fact that each of the interested parties looks respectively at their own interests, it is clear that each of them gives different weights to the categories. There is also another variable factor, namely the sector in which the company operates. For example, it is not possible to give the same weights to the categories for a company that operates in the advertising or social sector, and to one who works in the construction sector, or to an IT company with a marketing company. For this reason, the weights given to the categories will not change only for the various stakeholders but also for the sector in which the company in question operates. Having said that, the best solution is to create a series of tables as different models for business sectors, containing the four types of stakeholders analyzed.

Table 6.2 is an example of a model for the engineering sector in which weights for the different categories and the various stakeholders are shown. As this is a reputation system for evaluation of the degree of reliability of a target, some weights are suggested but it is advised that a more refined weight should be determined for each category.

	Business	Legal	Financial	Digital	HR	Social
M&A	30%	30%	30%	0%	5%	5%
JV	20%	25%	25%	15%	5%	10%
Customer	10%	35%	45%	5%	0%	5%
Supplier	20%	30%	20%	10%	20%	0%

Table 6.2: Weights for Categories in Reputation Index Model

In suggesting weights to apply to the various categories of the reputation index model, considering the engineering sector, different weights were applied for the different stakeholders.

In more detail, for a M&A candidate, the effect of the Business, Legal and Financial categories are considered to be of primary importance, therefore 30 percent has been given as the weight for each of them. The management of the company, the presence of ethical policies or environmental rules, and the presence of internal accidents are fundamental, as well as the legal aspects in which obviously no dispute with the law or the state must take place. Moreover, the financial aspects must be checked and the financial statements reviewed. All of these reasons give a high weight to these categories. A lower weight has been given to HR and Social categories since in a M&A of the sector specified before, the buyer isn't concerned about these fields. Moreover, no weight has been given to the Digital category because with an M&A both the site and the staff emails will no longer be used.

A JV candidate has roughly the same weights for the categories as of a M&A candidate, in fact the Business, Legal and Financial categories have a higher importance than the others. However, in this case more importance has been given to the Digital and Social categories seeing as having good feedback from customers or a good score on the digital measures affects the opinion of company that is looking to a JV candidate.

In the case of a Customer candidate, the Financial category is the one considered to be of primary importance. Knowing that the client has excellent balance sheets and has no problems paying for the products or services sold to him is of fundamental interest. The weight applied to this category is of 45 percent. Another substantial part

of the weight, 35 percent, was given to the Legal category, which will ensure that the customer is on good terms with the state and the law. On the other hand, very low weights have been given to the Digital and Social categories because in a Customer selection these categories are not relevant.

As far as the selection of suppliers is concerned, the weights given for each category are more evenly distributed. More importance has been given to the Legal category in which all the measures are important, from the presence of insurance policies, to the presence of patent and trademarks, and also checking that the supplier is on good terms with the state and the law. Business and Financial categories are important as well, but notice in this case that the same weight was given to the HR category. The HR category is relevant because during the selection of suppliers the status of the employees is evaluated as well as any disputes regarding them.

The weights given to the categories for the various stakeholders, even if motivated, are subjective and in agreement with the Alfa Group Company. Depending on the sector in which the target company operates, a different model will be used by the system. In a condition where the models created are not satisfactory for which they use the system, it is possible to create on the user interface of the system a section containing a form in which to insert the weights for a more dynamic evaluation of the reputation score.

6.3 Algorithm Implementation on the System

The measures to assess reputation have been defined through the index model, as well as the weights attributed to the various categories for the various stakeholders. This section will now analyse how this set of measures and weights was organized in the form of an algorithm implemented in the reputation system created.

Once the final report containing the information regarding the chosen target has been created, the system will start to apply the measures described in table 6.1, giving each measurement a score between 1 and 10. For example, in the case of the Digital category, the system will check the presence of any breach. In the affirmative case a count will be made and a score from 1 to 10 will be given on average to how many breaches were found. The system will perform this operation for each measurement in the table, then calculate an average of the individual measures for each category. Everything can be written using equation 6.1:

$$C_k = \frac{\sum_{i=1}^{n_k} x_{ki}}{n_k} \quad (6.1)$$

where

- C_k = average of measures for the category k
- x_{ki} = effective value of the category k for measures i
- n_k = total number of measures for the category k

The system now can give an overall score for each stakeholder, in fact, once the sector in which the target company works is found, a model similar to the one seen in the table 6.2 will be chosen. Applying the weights defined in the model, as a sort of

weighted average, the system will calculate the reputational score for each stakeholder. This is shown in equation 6.2

$$RS_z = \sum_{k=1}^{n_k} (C_k * p_{zk}) \quad (6.2)$$

where

- RS_z = effective reputational score of the stakeholder z
- C_k = value calculated through the equation 6.1
- p_{zk} = effective weight of the stakeholder z for the category k
- n_k = total number of categories k for the stakeholder z

Since the scale assigned to the measures on the table 6.1 is from 1 to 10, the reputational score calculated for each stakeholder through the previous formula will be included on the same scale. The higher the reputational score, the higher the reliability of the target as a candidate for that stakeholder.

The recursive search made to find the people or companies connected to the target described in the section 5.5, has brought to light that the reputation of the target can change based on the people or companies connected to it. In calculating the reputational score for each stakeholder this factor was not included because another series of measures that assesses the degree of reputation of the people or companies connected to the target should be defined, given that their connection to the target must be assessed properly.

The reputational score will be accompanied by a reliability score related to the score itself. As described in chapter 4.4, some information regarding the chosen target can be entered into the system through the user interface forms shown in chapter 5.6. For this reason, it was decided in agreement with the Alfa Group Company to give a different degree of reliability to this information than to that found through tools or services. The defined percentages are:

- 100% = degree of reliability of information found through OSINT tools and payable services
- 70% = degree of reliability of information received from third parties
- 50% = degree of reliability of information received from the target company or person

Once the percentages are defined, the calculation of the degree of reliability is very similar to the equations described above. For each measurement in table 6.1, an average of the percentages will be made between all the measurements in the table. This is shown in the equation 6.3

$$DoR = \frac{\sum_{i=1}^n x_i}{n} \quad (6.3)$$

where

- DoR = degree of reliability of the reputational score
- x_i = degree of reliability of measure i

The final score of the degree of reliability is unique for the entire report and therefore the same for each reputational score of the stakeholders.

Summary

The algorithm that provides the reputation score of the target according to the described stakeholders has been defined. The reputational score in fact indicates the reliability of the target company or person with a VAT number. The higher the degree of reputation, the more reliable the target will be as a candidate for the stakeholder.

Since the descriptions noted with the scale of 10 measures in table 6.1 are ideal, it is not anticipated that many companies would achieve a score of 10 for many of the measures of reputation. Thus, it is unlikely that many companies would be classified as the highest score. On the other hand few companies should fall into the lowest range of the scale. So the majority of companies should fall within the 8 to 6 score rating.

The reputational score will be accompanied by a degree of reliability of the score itself. Since some information regarding the target can be found by third parties or thanks to the company itself, their reliability must be considered, and therefore a reliability score for the reputational score has been calculated. Also in this case, the higher the degree of reliability, the greater the correctness of the reputational score.

7. Conclusions

The methodology used during the entire project was appropriate to the nature of the Alfa Group Company, and to the business order it was developing at that moment. To design and develop a system able to assess the reputation and the reliability of a company quickly and easily has always been the main goal of the entire project.

The purpose of this thesis is to develop a system able to assess the reliability of a target through a reputational algorithm that reflects differences in the meaning of reputation from perceptions through various stakeholder segments. The target analysed by the system will in fact be considered as a possible candidate for the defined stakeholders and therefore a different reputation score will be attributed for each of them. In this chapter, a summary of the entire project is first presented, followed by an evaluation and discussion of it. The findings are then discussed in terms of their implications for management. Finally, a comment on the company that allowed all this to be carried out and areas for future research on the topic are outlined.

The literature review focused on perceptions and evaluations of corporate reputation and due diligence, globally and by academics, in order to identify the context of the project. Using the literature review a reputation index model, containing all the key components and parameters necessary to define the reputation from an objective point of view, was then built. It needed to be rigorous and sensitive towards the different meanings that the concept of reputation has for various stakeholders. In fact, the literature review has also brought to light the different meanings of reputation and its components for the various stakeholders.

The initial stages of the project soon made apparent the importance of defining sources in which the information needed in order to assess the reputation of a target can be recovered. The opportunity to find information around the web through the OSINT tools and payable services, has allowed a system, which is able to assess the reliability of a target quickly and easily, to be created. In fact, the system, once the gathering data phase has been carried out, will process the information and then create a report which is accessible to the user.

The need to create a reputation evaluation algorithm is highlighted not only by the fact that the considered stakeholders evaluate reputation in different ways, but also by the fact that the report should show the degree of reliability of the target without having to read the report itself. For these reasons a reputational score, visible on the top of the report, has been given for each stakeholder involved. These scores allow the person or company that reads the report to understand immediately either to which role the target is more appropriate or whether it is a suitable candidate at all.

The report generated by the system will be of fundamental importance for a company or an analyst who will evaluate the position of the target as a candidate for one of the stakeholders previously seen. Thus, the reputation score will be important at first sight,

but then all the data contained in the report will be even more important. The report will be structured in the same way as the reputational index model, so the categories as well as the key components of the model will be highlighted. In addition, the report will contain a section dedicated to the people or companies connected to the target that were found during the recursive phase.

The created system is able to generate a report and evaluate it, thus showing the reliability of a chosen target. However, it should be emphasised that the information retrieved by the system may not be sufficient to evaluate the target. Therefore, the ability to insert information relating to the chosen target directly in the report was created. An investigation with the purpose of finding more information about a target can in fact be carried out by a team of people either who question the target or the parties connected to the target, such as suppliers, employees or customers, in order to find something that the system could not find. Information obtained through this investigation will increase or decrease the reputational score of the target but at the same time, it will not be possible to test their truthfulness. For this reason a degree of reliability has been given to the reputational score, showing how true or not it is.

Although the system created reflects exactly what was requested and thought of as a final solution by the company, various aspects were not considered and others were not developed to the full. Indeed, some considerations are necessary.

The system was provided with a graphical interface where information retrieved through the investigation phase previously discussed is inserted. The form present in the user interface has already been described during the previous chapters, but not the way the system should process the information entered. The form has sections that structure the information entered into the system by categories. The question is, how does the system evaluate the information entered for the reputation evaluation algorithm? We know that the algorithm is built on the basis of some measures that check the presence of a given element. For this reason the information that should be inserted into the form must already be elements that have a complete meaning for the system. In this context, once the investigation phase is completed, a careful analysis of the information followed by a classification is necessary before the data can be entered into the system. By performing these actions, the reputation algorithm will therefore be able to provide a more accurate reputational score for each stakeholder.

Another consideration, regarding the list of people and companies connected to the target, must be made. It has been mentioned during the project that this list influences the reputational score of the target. Inputting into the algorithm another measure that is considered during the final calculation of the reputation score, is a completely feasible thing, but how can this list of people and companies be evaluated? Can each element of the list be evaluated according to the same measures previously defined? Through these questions the problem is broad, in fact, a big discussion could be opened in this context. Taking for granted that the defined measures are suitable to be applied also to the companies connected to the target, other evaluation measures must be defined for people. Moreover, once the measurements have been defined, the respective weights must also be assigned. Due to time constraints, none of the topics mentioned above have been studied or defined even though it would certainly be interesting to explore them deeper.

The last consideration that must be made regards to the social category of the system is described below. Being aware of how the reputation of social networks influence thinking nowadays, the key component “Social Network” has been included in the reputation index model. The system in fact had to be equipped with an additional

service (SocialNet), as discussed in previous chapters, which analyses the main social networks of the target. For reasons of time and licenses, this payable service was not used and therefore no measure related to this key component was defined for the reputation evaluation algorithm. Having said this, it would undoubtedly be interesting and useful in improving the algorithm, to implement this service with the relative measures for the evaluation of the reputation.

7.1 Future Development

This project focused on assessing the reputation of a target for business purposes. A reputational score is given to the target based on a reputational model in order to assess its predisposition as a candidate to defined stakeholders. Implications for future research can be explored along the following pathways: Firstly, the application of the measurement tool to other business sectors and to other markets. Secondly, the creation of a database connected to the system in order to archive the generated reports. Thirdly, given the importance of the reputational score to understand the degree of suitability of the target as a candidate for the stakeholders, the system can be involved in a process that monitors a target or a range of targets in order to see how its reputational score varies with time.

There are opportunities not only to apply the reputation concept in different ways and in different business sectors but also to extend the reputational measurement approach towards organisations such as political parties, non-profit organisations and NGOs. Given the development of the reputation index model, once a company or organisation's stakeholders have been identified, a research could be conducted among the new stakeholder segments and relevant components, constituents and items identified in order to make a specific reputational index model for that sector.

Creating a database connected to the system in order to archive the generated reports could be a good solution for each company or person that is going to use the system often. The database could be structured with appropriate tables containing the report information. Each table with the related fields would correspond exactly to the categories and key components of the report, already defined in the reputational index model. There are various advantages of this structure: having the report already organised ensures that the structure is easier to create, the system can be connected directly to the database and data can be inserted directly, and there is the possibility to query the database easily and efficiently.

Another future research in which the system may be involved, is the creation of an automated process that constantly monitors a target or list of targets for a period of time in order to see their reputational score vary over time. In fact, the system can be integrated into a process where a request is automatically generated every N day to generate a report on a specific target. This future development could be associated with the previously mentioned one. In this context, however, the use of a database would be even more useful than it was before. Monitoring the reputation score of a target over time means controlling the contents in the various categories and key components of the reputation. Knowing how the information and the reputational score of the target change over time, would help in making a choice about the target as a candidate for a stakeholder and also create a hypothesis about its future performance.

7.2 The Alfa Group Company

The Alfa Group Company was strongly cited throughout the project developed, in fact, it was not limited only to providing a project for the creation of the thesis, but also gave a considerable contribution throughout the development of the project.

A description of the company is essential. Alfa Group is a software house that has been operating in the market of services and solutions since 1996 and is one of the most dynamic and innovative in the ICT sector, specialising in solutions for enterprise management and security. The workforce is made up of staff with specific expertise in areas of interest to the company and it has one hundred employees and fifty consultants at present in Italy, with headquarters in Rome, Milan, Bologna and Fermo, and in Utrecht, Holland. The Fermo site, in which I carried out my master's thesis, specialises in the conception, design and implementation of IT solutions. From those concerning the development of management systems for small and medium-sized enterprises to those concerning cyber security and those concerning business processes management.

In agreement with the University of Camerino, the Alfa Group Company made a proposal regarding the development of the "Reputation System" project which was already being worked on. They offered me the possibility to develop the project internally alongside an employee who guided me throughout the development. The help from the company, especially from the managers Gaetano Ascenzi and Andrea Carassai, led to a different approach in the project. From the methodology used, to the decisions made regarding which tools to use or which stakeholders to involve, without forgetting the technical support received.

The success of the project was greatly influenced by the presence of the company and their interest in helping me to complete the thesis. The Alfa Group Company has been a source of growth, both technically and personally. Thanks to it I deepened my reputation culture within the company, I learned new services and tools, and I understood the internal dynamics of a company. It has been a very positive experience that will help me in the world of employment.

A. OSINT tools output Pictures

This appendix contains pictures related to the output of OSINT tools implemented in the system. The domain *unicam.it* was used as an example.

The picture A.1 shows a list of email addresses retrieved through requests made to different data sources during the OSINT research phase. Most of them actually exist but others are incorrect. During the parsing phase, in fact, some patterns may not recognise the text, and therefore an incorrect email address will be returned.

The pictures A.2 and A.3 show a list of hostnames retrieved through requests made to different data sources during the OSINT search phase. The pictures also show the IP address, country, region, latitude and longitude of each host name. Similar to the issue regarding the email addresses, some hosts may be incorrect or non-existent, this is always due to patterns applied to the text.

michele.savini@studenti.unicam.it
luca.tesei@unicam.it
<valerio.passini@unicam.it
address:giovanni.caprioli@unicam.it
augusto.amici@unicam.it
riccardo.pettinari@unicam.it
eleonora.paris@unicam.it
daniele.tomassoni@unicam.it
nisha.dahiya@unicam.it
sgi@pec.unicam.it
delegato.tutorato@unicam.it
idea4b@unicam.it
tutorato@unicam.it
matematica@unicam.it
radio@unicam.it
infocs@unicam.it
luca.petrelli@unicam.it
risoremarinecostiere@unicam.it
0737402540david.vitali@unicam.it
Cingolani0737402126segreteria.scienze@unicam.it
merchandising@unicam.it
cinfo@unicam.it
barbara.mastrocola@unicam.it
sst@pec.unicam.it
biblioteca.sad@unicam.it
marco.bello@unicam.it
luciano.barboni@unicam.it
segreteriastudenti.scienze@unicam.it
annamaria.santroni@unicam.it
pensy@unicam.it
rettore@unicam.it
protocollo@pec.unicam.it
roberta.tacchi@unicam.it
ortobotanico@unicam.it
giovanna.sabbieti@unicam.it
polomuseale@unicam.it
gianluca.marucci@unicam.it
giulio.tomassini@unicam.it
synbiotec@unicam.it
pierluigi.palombi@unicam.it
marco.antonini@unicam.it
dirittocivile@unicam.it
elide.ferranti@unicam.it
andrea.catorci@unicam.it
ludovico.romagni@unicam.it
barbara.mastrocola@unicam.it
orientamento@unicam.it
loredana.cappellacci@unicam.it
ilo@unicam.it
anna.silano@unicam.it
lucia.ruggeri@unicam.it
placement@unicam.it
scuola.dirittocivile@unicam.it
stefania.pucciarelli@unicam.it
manuela.prenna@unicam.it
facolta.veterinaria@unicam.it
francesco.nobili@unicam.it
bernardino.gentili@unicam.it
roberto.ballini@unicam.it
gabriella.gabrielli@unicam.it
francesca.magni@unicam.it
emanuela.merelli@unicam.it
<emanuela.merelli@unicam.it
<luca.tesei@unicam.it
luca.tesei@unicam.it
paolo.polidori@unicam.it
cuia.presidenza@unicam.it
mirko.mancinelli@unicam.it
carlo.santini@unicam.it
mario.cocchioni@unicam.it
cristiana.pascucci@unicam.it

Figure A.1: Email addresses retrieved by OSINT tools implemented on the system

Hostname	IP Address	Country	Region	Latitude	Longitude
root.camcic.unicam.it	root.camcic.unicam.it	Italy	Toscana	43.1399	43.1399
www.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
www.cs.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
docenti.unicam.it	193.204.8.131	Italy	Marche	43.1399	13.0692
d7.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
gnxas.unicam.it	193.205.94.20	Italy	Marche	43.1399	13.0692
arwin.unicam.it					
darwin.unicam.it					
studenti.unicam.it					
superfici.unicam.it					
cms.unicam.it	193.204.8.48	Italy	Marche	43.1399	13.0692
disabili.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
ms.unicam.it					
saad.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
eb.unicam.it					
web.unicam.it	193.204.8.30	Italy	Marche	43.1399	13.0692
idattica.unicam.it	idattica.unicam.it			0	0
didattica.unicam.it	130.186.27.50	Italy	Emilia-Romagna	44.4756	11.2759
international.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
edea.unicam.it	edea.unicam.it			0	0
medea.unicam.it	193.204.8.31	Italy	Marche	43.1399	13.0692
iblioteche.unicam.it					
camcic.unicam.it	193.204.8.13	Italy	Marche	43.1399	13.0692
biblioteche.unicam.it					
web2.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
alwks.unicam.it					
vele.unicam.it					
isas.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
ail.unicam.it	ail.unicam.it			0	0
master.unicam.it	193.205.92.67	Italy	Marche	43.1399	13.0692
mail.unicam.it	mail.unicam.it			0	0
df.unicam.it	193.205.94.12	Italy	Marche	43.1399	13.0692
f.unicam.it	f.unicam.it			0	0
dns.df.unicam.it	193.205.94.11	Italy	Marche	43.1399	13.0692
omputerscience.unicam.it	omputerscience.unicam.it			0	0
computerscience.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
learning.unicam.it	learning.unicam.it			0	0

Figure A.2: List of hostnames retrieved by OSINT tools (first)

Hostname	IP Address	Country	Region	Latitude	Longitude
masterenergie.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
formorientamento.unicam.it					
labtec.unicam.it	94.177.192.179	Italy	Toscana	43.4614	11.8769
masterwineexportmanagement.unicam.it					
gnxas.unicam.it	193.205.94.20	Italy	Marche	43.1399	13.0692
bias.unicam.it					
scienzeambientali.unicam.it					
masternutrizione.unicam.it	193.204.14.15	Italy	Marche	43.1399	13.0692
infm.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
disabili.unicam.it					
studenti.unicam.it					
masterisorsemarinecoziere.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
farmacologia.unicam.it	193.204.8.30	Italy	Marche	43.1399	13.0692
www.tavolaperiodica.unicam.it	193.204.8.30	Italy	Marche	43.1399	13.0692
tavolaperiodica.unicam.it	tavolaperiodica.unicam.it			0	0
erasmus.unicam.it					
tiezzi.unicam.it	172.217.14.83	Netherlands	Noord-Holland	52.374	4.88969
archspace.unicam.it					
masterorienta.unicam.it					
convegni.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
nle.unicam.it					
chimali2018.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
bioetica.unicam.it	193.204.8.30	Italy	Marche	43.1399	13.0692
cs-research.unicam.it	64.233.177.121	United States	California	37.406	-122.079
orientamento.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
sbmv.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
entervista.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
www.cs.unicam.it	94.177.192.171	Italy	Toscana	43.4614	11.8769
cs.unicam.it					
portal.unicam.it	193.204.8.3	Italy	Marche	43.1399	13.0692

Figure A.3: List of hostnames retrieved by OSINT tools (second)

B. User Interface Pictures

This appendix contains the pictures related to the user interface of the created system.

The picture B.1 shows the user interface home page where there is a form in which to insert the target details.

The picture B.2 shows the user interface company section in which a big form is presented. It can be used to insert the information retrieved by the investigation phase from the target company.

The picture B.3 shows the user interface network of people section in which a tree graph is presented. The tree graph shows the people or companies connected to the target, every node is marked with its VAT or VAX code and each edge is marked with the role that the node has.

The picture B.4 shows the user interface report section in which, once the report is successfully created, you can open it with the appropriate button.

REPUTATION SYSTEM



Business Report form

Company Name

VAT or TAX code

Company Website

Search

Figure B.1: User Interface home page



Data Received from the Company

Business	Financial	Legal
R&D Strategy		
Ethical Balance Sheet		
Liability claims		
Joint Venture Agreements		
Employees Satisfaction		

Search

Figure B.2: User Interface company section

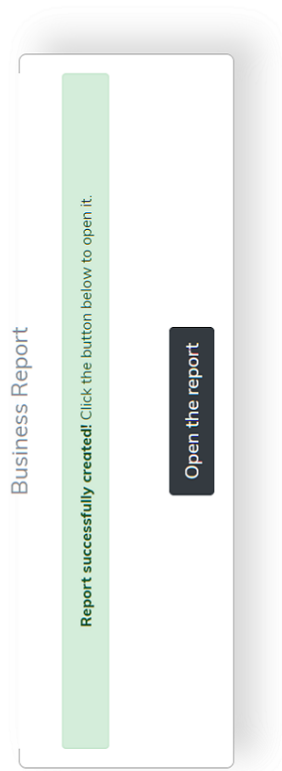


Figure B.4: User Interface report section

C. Report Pictures

This appendix contains the pictures related to an example of the report generated by the system. They show how the report should look and how it is structured.

The picture C.1 shows the first page of the report in which the main target details and the reputational score for each stakeholder are underlined at the top of the page. The balance sheets and the relevant event fields are present in this page as well.

The picture C.2 shows the second page of the report in which the other fields are presented, from the ones containing the legal data to the ones containing the business administration people, and the ones containing the list of people connected to the target.



Reputation System Report

Company Name

Summary	Reputational Score	
Registered Office [REDACTED] City Seat [REDACTED] VAT Code [REDACTED] Status Activity [REDACTED] VAX Code [REDACTED] ATECO 2007 [REDACTED] Date of Incorporation [REDACTED] Business start Date [REDACTED] Company in [REDACTED] Legal Nature [REDACTED] Phone number [REDACTED] Phone number [REDACTED] Certified email [REDACTED]		
	Score	Reliability
	M&A	[REDACTED]
	JV	[REDACTED]
	Customer	[REDACTED]
	Supplier	[REDACTED]

Balance summary

	Year xxxx	Year xxxx	Year xxxx
Economic Data			
Production value	[REDACTED]	[REDACTED]	[REDACTED]
Production costs	[REDACTED]	[REDACTED]	[REDACTED]
Financial incomes	[REDACTED]	[REDACTED]	[REDACTED]
Profit / Loss	[REDACTED]	[REDACTED]	[REDACTED]
Cash Flow	[REDACTED]	[REDACTED]	[REDACTED]
Financial Data			
Total credits	[REDACTED]	[REDACTED]	[REDACTED]
Total debts	[REDACTED]	[REDACTED]	[REDACTED]
Net Assets	[REDACTED]	[REDACTED]	[REDACTED]

	Year xxxx	Year xxxx	Year xxxx
Balance sheets			
ROI	[REDACTED]	[REDACTED]	[REDACTED]
ROE	[REDACTED]	[REDACTED]	[REDACTED]
ROS	[REDACTED]	[REDACTED]	[REDACTED]

Relevant Events

[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]

Figure C.1: Example Report (first page)

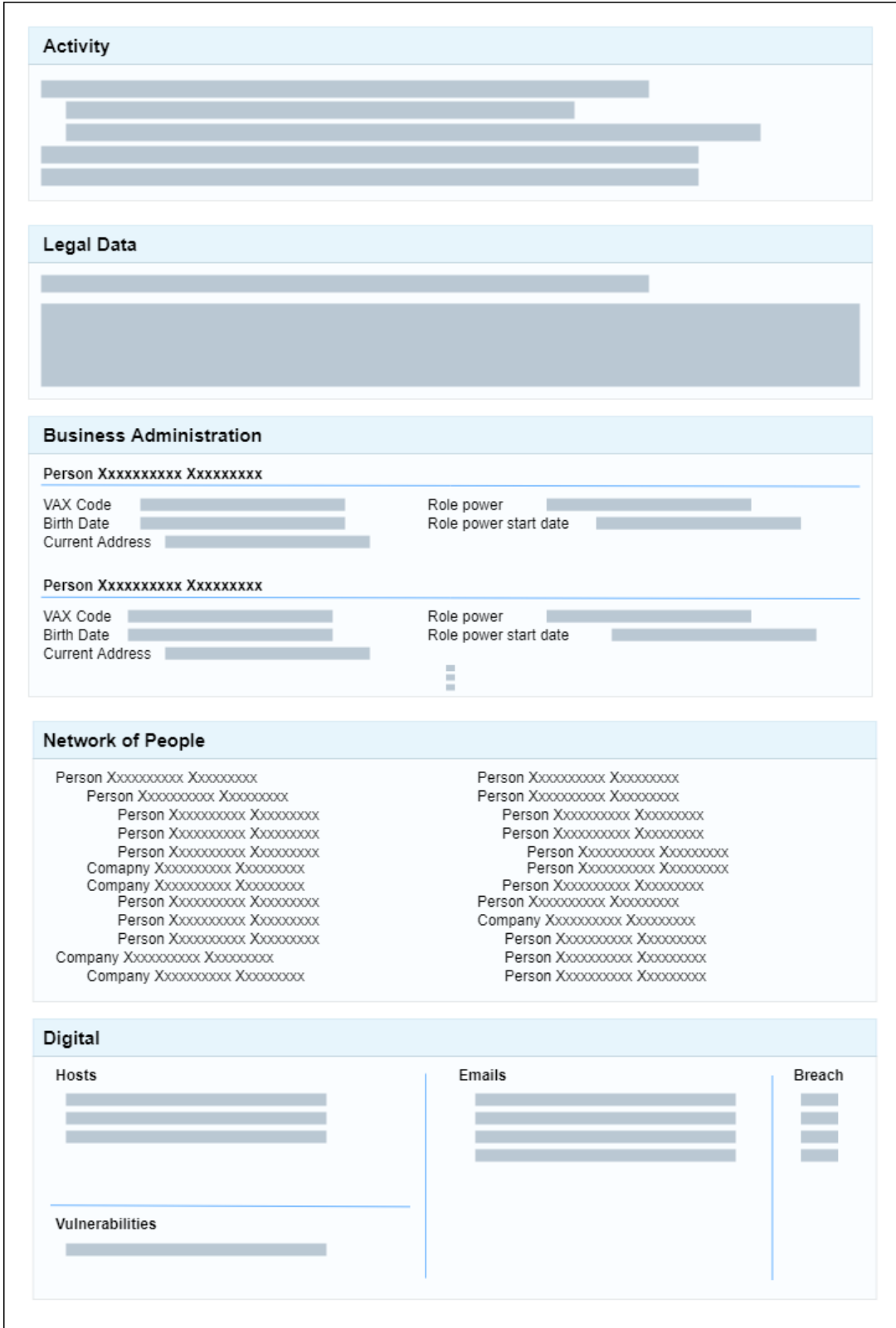


Figure C.2: Example Report (second page)

Bibliography

- [1] S. Balaji and M. Sundararajan Murugaiyan. “Waterfall Vs V-Model Vs Agile: A Comparative study on SDLC”. In: *International Journal of Information Technology and Business Management* 2.1 (2012), pp. 26–30.
- [2] Weigelt K. and Camerer C. “Reputation and corporate strategy: A review of recent theory and applications”. In: *Strategic Management Journal* 9.5 (1988), pp. 443–454.
- [3] Fombrun C. and Shanley M. “What’s in a name? Reputation building and corporate strategy”. In: *Academy of Management Journal* 33.2 (1990), pp. 233–258.
- [4] Charles J. Fombrun. “Reputation: Realizing Value from the Corporate Image”. In: Harvard Business School Press, Boston, 1996.
- [5] Fombrun C. and Van Riel C. “The reputational landscape”. In: *Corporate Reputation Review* 1 (1997), pp. 5–13.
- [6] Cable D. M. and Graham M. E. “The determinants of job seekers’ reputation perceptions”. In: *Journal of Organizational Behavior* 21.8 (2000), pp. 929–947.
- [7] Deephouse D. L. “Media reputation as a strategic resource: An integration of mass communication and resource-based theories”. In: *Journal of Management* 26.6 (2000), pp. 1091–1112.
- [8] Bromley D. B. “Relationships between personal and corporate reputation”. In: *European Journal of Marketing* 35.3/4 (2001), pp. 316–334.
- [9] Mahon J. F. “Corporate reputation: A research agenda using strategy and stakeholder literature”. In: *Business and Society* 41.4 (2002), pp. 415–445.
- [10] Whetten D. A. and Mackey A. “A social actor conception of organizational identity and its implications for the study of organizational reputation”. In: *Business and Society* 41.4 (2002), pp. 393–414.
- [11] Petkova A. P. Rindova V. P. Williamson I. O. and Sever J. M. “Being good or being known: An empirical examination of the dimensions, antecedents, and consequences of organizational reputation”. In: *Academy of Management Journal* 48.6 (2005), pp. 1033–1049.
- [12] Rhee M. and Haunschild P. R. “The liability of good reputation: A study of product recalls in the U.S. automobile industry”. In: *Organization Science* 17.1 (2006), pp. 101–117.
- [13] Carter S. M. “The interaction of top management group, stakeholder, and situational factors on certain corporate reputation management activities”. In: *Journal of Management Studies* 43.5 (2006), pp. 1145–1176.
- [14] Jermier J. M. Barnett M. L. and Lafferty B. A. “Corporate reputation: The definitional landscape”. In: *Corporate Reputation Review* 9.1 (2006), pp. 26–38.

- [15] Kent Walker. “A Systematic Review of the Corporate Reputation Literature: Definition, Measurement, and Theory”. In: 12.4 (2010), pp. 357–387.
- [16] Fombrun Charles J. “Corporate Reputation – Its Measurement and Management”. In: *Thesis* 4 (2001), pp. 23–26.
- [17] John M.T. Gray Edmund R./Ballmer. “Managing Corporate Image and Corporate Reputation”. In: *Long Range Planning* 31 (1998), pp. 695–702.
- [18] Richard Hall. “The Strategic Analysis of Intangible Resources”. In: *Strategic Management Journal* 13 (1992), pp. 135–144.
- [19] Manfred Schwaiger. “Components and Parameters of Corporate Reputation - an Empirical Study”. In: *Schmalenbach Business Review* 56 (2004), pp. 46–71.
- [20] C. Eidson and M. Master. “Top ten...most admired...most respected: Who makes the call?” In: *Across the Board* 37.1 (200), pp. 16–22.
- [21] Joy M. Sever Charles J. Fombrun Naomi A. Gardberg. “The Reputation QuotientSM: A multi-stakeholder measure of corporate reputation”. In: *The Journal of Brand Management* 7.4 (2000), pp. 241–255.
- [22] Gerald E. Fryxell. “The Fortune Corporate ‘Reputation’ Index: Reputation for What?” In: *Journal of Management* 20.1 (1994), pp. 1–14.
- [23] Linda S. Spedding. *Due Diligence Handbook. Corporate Governance, Risk Management and Business Planning*. Elsevier, 2009.
- [24] Robert F. Lusch Michael G. Harvey. “Expanding the nature and scope of due diligence”. In: *Journal of Business Venturing* 10 (1995), pp. 5–21.
- [25] C. Gates. “Get the most from your acquisition”. In: *Journal of Business Strategy* 9.6 (1988), pp. 8–11.
- [26] Kroener P. and Kroener W. “Towards more successful mergers and acquisitions”. In: *International Journal of Technology* 6.2 (1991), pp. 33–40.
- [27] Heame J. and Dean D. “The preacquisition search for environmental risks”. In: *Mergers and Acquisitions* (1989), pp. 33–39.
- [28] P. Kotler. “From sales obsession to marketing effectiveness”. In: *Harvard Business Review* (1977), pp. 67–75.
- [29] Mallin W. Garelik M. and Peterson J. “Smart HR planning boosts merger success”. In: *Savings Institutions* (1992), pp. 26–29.
- [30] Sridhar Ramamoorti Karen Cravens Elizabeth Goad Oliver. “The Reputation Index: Measuring and Managing Corporate Reputation”. In: *European Management Journal* 21.2 (2003), pp. 201–212.
- [31] Consumers International. “The State of Data Protection Rules Around the World, A Briefing for Consumer Organisations”. In: *Consumers International, Coming together for change* (2018).
- [32] Natasha Singer Mark Scott. “How Europe Protects Your Online Data Differently Than the U.S.” In: *The New York Times* (2016).
- [33] Rami Hijazi Nihad A. Hassan. *Open Source Intelligence Methods and Tools. A Practical Guide to Online Intelligence*. Apress, 2018.
- [34] Tim Tomes. *Recon-ng*. URL: <https://bitbucket.org/LaNMaSteR53/recon-ng>.

- [35] Christian Martorella. *theHarvester*. URL: <https://github.com/laramies/theHarvester>.
- [36] registroimprese.it. *registroimprese.it, I dati ufficiali delle Camere di Commercio*. URL: <https://www.registroimprese.it>.
- [37] Cribis company and dunbradstreet. *Cribis, a CRIF company*. URL: <https://www.cribis.com/>.
- [38] ShadowDragon. *SocialNet*. URL: <https://shadowdragon.io/product/socialnet/>.
- [39] Bill Burke. *RESTful Java with JAX-RS 2.0. Designing and Developing distributed Web Services*. O'Reilly Media, Inc., 2014.
- [40] W3C. *Web Services Description Language (WSDL)*. 2001. URL: <https://www.w3.org/TR/wsdl.html>.
- [41] Henry S. Thompson Shudi (Sandy) Gao C. M. Sperberg-McQueen. *XML Schema Definition Language*. 2012. URL: <https://www.w3.org/TR/xmlschema11-1/>.