UNIVERSITÀ POLITECNICA DELLE MARCHE

Faculty of Engineering

Department of Information Engineering

Master of Science in Biomedical Engineering

TOOLS FOR THE ASSESSMENT OF CYBER RISK IN HEALTHCARE FACILITIES

Supervisor:

Prof. Marco Baldi

Candidate:

Rosita Laratta

Co-Supervisors:

Prof. Franco Chiaraluce

Giulia Rafaiani

Abstract

Nowadays, because of the spreading of healthcare technologies new apprehensions arise, especially concerns cyber-security within healthcare. As a result, engineers and regulators need to expand pre-existing risk management frameworks tailored to safety to also encompass cybersecurity. As the European Data Protection Regulation (GDPR 679/2016) comes into force, organizations and in particular healthcare companies are enforced to adopt advanced measures to address cyber-risks. The idea of a risk-based approach is highly recommended by the Regulation and is inspired by ISO 31000 (Risk management). The pre-existing risk management approaches are classified into quantitative, qualitative, and semi-quantitative ones; the choice is strictly related to the context, purpose, and objectives of the model or experiment proposed. A Risk Management framework aims to enable every organization to identify, prevent and manage all the impending risks within its business and fully illustrate the delicate and fundamental process of risk management through a structured approach. In 2015 the National Framework for Cybersecurity was presented, inspired by the Cybersecurity Framework created by NIST (National Institute of Standards and Technology) as an adequate support tool for organizations that need strategies and processes aimed at protecting personal data and cyber-security. The aim is principally the reduction of risks linked to cyber-threats. One of the mandatory requirements of the Regulation is the formation of the data processing activities, as one of the main tools to accomplish accountability. The Data Controller, through the records of data processing activities, should demonstrate compliance with GDPR in a proportional way to the risk associated to protect personal data. To create a new tool, we collaborated with Azienda Ospedaliero Universitaria Ospedali Riuniti for the assessment of cyber-risk. The current study aims to accomplish the complex requirements of the Regulation thanks to a new tool proposed in electronic format and to perform a risk assessment for any data processing activities. After proper evaluation, the most critical part of the data processing activities has been the analysis of technological context and IT Infrastructure. As a result, a risk index assessment could not be evaluated because of the limitations possessed. Therefore, we proposed a new approach to overcome the assessment methodologist problems

previously pointed out, combining the evaluation of the maturity index of the processes inherent to firms along with the evaluation of the complexity index of the desired firms with chosen infrastructure concluding with the attractiveness of the organization. Unfortunately, due to the poorness of information and case studies present in the literature, this goal has not been achieved, even though a parameterization has been implemented.

Index

1 Introduction

This increasingly modern new world has pointed out several important and severe developments raising new challenges for the governance and maintenance of public safety. Safety best practices are well exploited and utilized in critical sectors such as healthcare, transport, aerospace, and energy, and are likewise well established and prescribed by safety standards methods. Such standards stipulate the manner under which systems should be developed, verified, and maintained to minimize the risks (present and future) of accidents and failure over their entire lifetime. Yet, the entire established subsets of safety practices fall short of addressing the cybersecurity threats that ensue from the growing interconnectivity of formerly isolated systems [**1**].

As systems become more and more vulnerable to remote attacks, protection is essential as a mandatory assessment from both accidents as well as malicious cyber incidents – the safety and security requirements of these systems converge. As a result, engineers, decision-makers, and regulators need to expand established and pre-existing, *Risk Management* frameworks, subsets of controls and best practices, security standards, and regulations tailored to safety to also encompass cyber-security [2].

2 Risk Management Frameworks

Several methods have been developed so far, being classified in *quantitative, qualitative, and semi-quantitative* approaches of evaluating risk. Clearly, each risk assessment approach has advantages and disadvantages [4].

When exploiting a *quantitative assessment*, the goal would be to calculate numeric values associated with each component that result after risk evaluation. To obtain such scope, this method usually employs a set of methods, principles, or rules and, of course, its result would be rigorous, repeatable, and reproducible. In addition, the estimation of the events' probabilities and impacts can be compared rigorously, directly and, objectively. The real value of the assets must be determined, taking into consideration the cost of replacement, the cost of the productivity loss, the cost of brand reputation damage and, other values that represent direct or indirect assets for the organization. Even after the estimation of probabilities and impacts, the results might not be clear to determined and interpret, since is a very challenging task, consequently interpretation and explanation are required [5]. In addition, the benefits might be outweighed by the costs (potential estimating loss in case of an attack also determined in terms of the expert time and effort) and the possible requirements of tools to make the assessments [3].

Shifting to *qualitative assessment*, it is easy to understand why the beforementioned approach Is mainly exploited by small organizations. It typically employs a set of methods, principles, or rules-based on non-numerical categories or levels for assessing the risk, so not based on statistical or exact values to estimate the risk in an organization, but it is mainly based on relative values being used as data entries for the calculation of potential loss. Qualitative assessment methods seem to be time and cost-efficient since no statistical effective values are calculated to validate the model. This typology of assessment might be exploited to estimate the risk, losses, possible outcomes in an approximative mode, to easily categorize possible areas of improvement. Concluding, it is important to assess that the evaluation of the risk and its result are subjective, so the performance of *Risk Management* are hard to follow [3] since the results almost depend on the quality of the

Risk Management team making comparisons troublesome. Another disadvantage of using qualitative assessment methods is that a cost-benefit analysis is not implemented, only a subjective approach of the author and that makes difficult the implementation of controls since they provide approximate evaluations for both likelihood and impact [3].

Finally, *semi-quantitative assessment* typically employs a set of methods, principles, or rules for assessing the risk that uses ranges of numbers, scales, or representative numbers. It usually provides an intermediary level between the evaluation of *qualitative risk assessment*, as textual, and the numerical evaluation of *quantitative risk method assessment*, providing a score at the end of the assessment to obtain the risk. Therefore, it produces more consistent and rigorous outcomes than qualitative assessment provides, also avoiding some greater ambiguities that might arise from the beforementioned approach. The ranges of numbers or the scales translate easily into qualitative terms, making easier the presentation of the results, but, at the same time, also allow relative comparisons between values in different ranges or even within the same range [3] However, the combination and interpretation of the results can be hard, because of different rating scales.

Once all the three methodologies have been explained, the advantages and disadvantages listed and explained, the successive step is to clarify three kinds of approaches commonly used in order to accomplish a fine *qualitative* or *quantitative method risk assessment*, depending on our model/experiment purpose.

1. Starting from the first one encountered, it implies the use of the most common standards that concern the world of information security (in particular ISO/IEC 27001) [7] and ISO/IEC 27017 [8] for the Cloud). In this very specific case, risk analysis is part of the *Information Security Risk Management* process which, according to [9] can be easily synthesized in its proceeding steps:

- ✓ **Establishing the context** defines the **scope** for the *Risk Management* process and **sets the criteria** against which the risks will be assessed. The scope should be determined within the context of the firm's organizational objectives. Risks are uncertainties that affect the

achievement of business objectives, so the risks cannot fully be identified if these objectives and strategies are unclear. Practically is the identification and classification of the assets chosen for the analysis, external and internal factors that may currently impact the firm [10].

- ✓ **Identification of the risks**: so, threats and vulnerabilities. If a potential risk is not identified at this stage it is omitted from further analysis, which means a material risk may be given insufficient attention. The risks that relate to the firm's context and business objectives must be identified, whether or not they are under the influence of the firm.

- ✓ **Analyzing and evaluating risk**: so, assessment criteria to assists in identifying, analyzing, and prioritizing key business risks. It helps validate and prioritize key risks to monitor and it highlights any opportunities for improvements to current activities used as controls in the business. A risk assessment provides insight into significant inherent risks from a practice perspective and links these to a firm's objectives, strategies, and business processes.

- ✓ **Developing a risk treatment plan:** risk treatment involves developing a range of options for mitigating the risk, assessing those options, and then preparing and implementing action plans. The highest-rated risks should be addressed as a matter of urgency.

- ✓ **Monitoring and reviewing**: the beforementioned should be a planned part of the *Risk Management* process and involve regular checking or surveillance. The results should be recorded and reported externally and internally, as appropriate. The results should also be an input to the review and continuous improvement of the firm's risk management framework.

A limitation of this approach derives from the threat/vulnerability/risk association; in a context in which risks are constantly evolving, it is not easy to identify an exhaustive list of threats, nor to assess the probability of their occurrence, especially if it is based on statistical evaluations of events

that have occurred in the past. It is necessary to ensure a periodic review and continuous improvement.

2. Shifting to the second approach, it is based on assessments and questionnaires, mostly audit, derived from best practices (the ISO/IEC families, the CIS Controls [11], NIST SP-800 [3], and on FNCS [12] and AgID [13] in Italy). It involves the definition and contextualization of the subsets of controls utilized in the reference frame chosen, through assessments and/or interviews, and, then, evaluating the level of implementation of these subsets of controls within the considered organization by calculating a maturity index for each homogeneous group of controls. This approach is mainly based on the assumption that a deviation from the best practice chosen as a reference corresponds to an increase in risk.

The beforementioned assumption might represent a major limitation regarding the effectiveness of the risk assessment method chosen, especially starting from the fact that the context of the organization and exogenous factors are not considered in the evaluation, only internal or intrinsic factors. Furthermore, the information collected through audits or interviews during the assessment might be difficult to interpret and objectivized since roles and responsibilities within the firms are not always clearly understood and defined leading to a lack of information collected or an impartially detailed outcome.

3. Last but not least, the third approach involves risk scenario simulation techniques exploiting ad utilizing international frameworks of the ISACA family (i.e. COBIT [14]). This approach stands as a powerful tool to describe the company risk scenario and the operational reality of ventures.

The exploitation of this kind of approach involving risk scenarios simulation techniques addresses problems related to ventures' maturity of organizational procedures and to the structure of the business models to which they refer. They are feasible and generate value only if combined with transverse IT governance protocols.

The next paragraph would amply explain *Risk Management* and the international reference standard that rely on upon under the aforesaid definition with peculiar attention to *National Cyber-security Frame.*

2.1 Risk Management and ISO 31000

The concept comprises all actions aimed at identifying and influencing the risks and opportunities arising from the actions and activities of a company and which could have positive or negative effects on the company itself [15]. On top of the definition of *ISO 31000:2018*, there is: "coordinated activities to direct and control an organization concerning risk" [9]. As the legislation continues, it is necessary to analyze and evaluate *ISO 31000* in a very punctual way some fundamental definitions and requirements [9]:

- ✓ ***Risk*****:** it usually relates cause and effect when talking of uncertainty on objectives. This definition is generally translated into the mathematical formula as the product between the likelihood of occurrence of an adverse event and the impact index that this event carries along his path.
- ✓ ***Effect*****:** the occurrence or change of a set of circumstances and it can be both positive and negative.
- ✓ ***Objectives*****:** is the object of the risk, can refer to different kind of activities and categories, according to the specific case.
- ✓ ***Control:*** the measure that maintains and/or modifies risk.

The task of *Risk Management* is not to eliminate all the risks, which would be practically impossible, but to process a series of applications, subsets of controls, guidelines that must be accomplished systematically.

It aims to enable every organization to identify, prevent and manage all the impending risks within its business and fully illustrate the delicate and fundamental process of *Risk Management* through a structured approach.

In general, *Risk Management* refers to the architecture (principles, frame of reference and process) for effectively managing the risks, while managing the risk refers to applying that architecture to specific risks. It is thought of as an iterative process and can be represented with the following scheme [9].

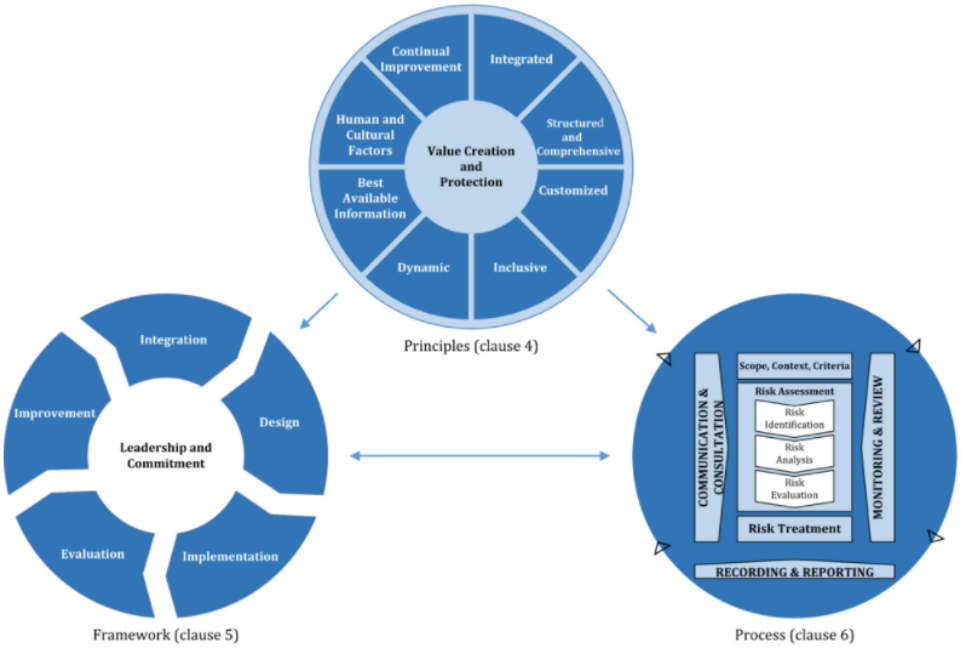

Figure 1: Principles, Framework, and Processes concerning Risk Management

Let's analyze some aspects and important definitions used in the process of *Risk Management*. The whole process consists of *a series of sequential* and *transversal phases* [16]. *The scope, the context,*

and criteria must be defined for customizing the *Risk Management* process, enabling effective risk assessment and suitable risk treatment *(1st sequential phase).*

- ✓ A proper *risk assessment* methodology is comprised of ***risk identification***, ***risk analysis*** **and** ***risk evaluation***. Risk assessment should be conducted systematically and *iteratively (2nd sequential phase).*

    i. ***Risk identification:*** consists of finding, recognizing, and describing risks in such a way that might help or prevent an organization achieving its goals.
    ii. ***Risk analysis:*** consists of comprehending the nature of risk and its features. This step is necessary for quantifying the risk. Risk analysis considers uncertainties, risk sources, consequences, likelihood, events, scenarios, controls, and their impact.
    iii. ***Risk evaluation:*** is necessary for supporting decision-making. It consists of comparing the results of the risk analysis with the defined risk criteria to determine where additional measures are required.

During *risk treatment* some actions take place as *transversal phases*:

- ✓ **Communication**: promotes awareness and understanding of the risk, whereas consultation involves feedback and information to help decision-making.
- ✓ **Monitoring and review** are needed to assure and improve the quality and effectiveness of the whole process.
- ✓ **Recording and reporting**: the *Risk Management* process and its outcomes through appropriate mechanisms are always suitable.

2.2 National Framework for Cybersecurity 2.0

Trying to address the topic of data protection is a difficult task since it requires organizations to equip themselves with suitable measures to protect the confidentiality, integrity, availability of

data, as well as guaranteeing the resilience of the processes when processing the aforementioned data.

In order to obtain an adequate support tool for organizations that need strategies and processes aimed at protecting personal data and cybersecurity in 2015 the National Framework for Cybersecurity was presented, inspired by the Cybersecurity Framework created by NIST (National Institute of Standards and Technology) [17]. The framework acts as a guide to increase the level of cybersecurity and data protection and, at the same time, contains recommendations on how to organize cyber-security *Risk Management* processes. It is predominantly created to support personal data protection strategies and cyber-security management. The aim is principally the reduction of the risks linked to the cyber threat. The approach of the proposed Framework is strictly based to a risk analysis, not linked to technology standards at all.

Even though the Italian National Framework for Cybersecurity is heavily inspired by NIST's Cybersecurity Framework it maintains a strong awareness of the context of the national landscape, with a specific focus on small and medium enterprises (SMEs), as it was predominantly designed for the particular Italian production context, independently from a specific production sector.

The National Framework for Cybersecurity is a common reference and has nothing to do with security standards as their aim differs inherently. A security standard tries to categorize existing or future standards and regulations while the endorsement to the National Framework is intended as it analyzes voluntary agreement.

As an enrichment of its informative references, the Framework 2.0 deals with: compliance with GDPR [6], D.Lgs. 65/2018 (transposition of NIS Directive) [3], and AgID [13].

The National Framework Core maintains the structure and division into 5 *functions*, 22 *categories*, 98 *sub-categories* and, *informative references*, like NIST's Cybersecurity Framework while introducing two new concepts: priority levels and maturity levels that manufacture the National Framework suitable for SMEs.

Each *sub-category* of the National Framework is not self-referential and represents a recommendation area, leaving to businesses the decision to implement by referring to the specific sector standard or regulation. Besides, the NIST Framework provides references to existing standards and Frameworks for each *sub-category*. Those references cover most of the pre-existing one already implemented by International organizations, such as the NIST Standard, the ISO/IEC and the COBIT Standards.

The Framework Core, Profile, and Implementation Tier of the Italian National Framework derives from the NIST Framework. Shortly, these concepts are going to be described.

The *Framework Core* is considered the *Core* structure of the management process of cyber-security as it comprises both technical and organizational points of view. It is structured hierarchically, as it has already been seen, into *Function, Category, and Sub-category*. The categories of concurrent and continuous *functions* are *Identify, Protect, Detect, Respond, Recover* and they represent the main topics to deal with to strategically obtain suitable Cyber-Risk Management. The Framework provides information about detailed resources, defining processes, and technologies to be put in place to manage the sole *Function, Category, and Subcategory*. Finally, the *Framework Core* structure shows informative reference, informative references that link the single *Sub-category* to several security practices by using sector standards (ISO, sp800-53r4, COBIT-5, SANS20 and others).

The 5 functions are briefly described below:

- ✓ *Identify*: it is linked to the understanding of the company context, of assets that support the critical business processes and relevant associated risks. The *Categories* within this Function are: *Asset Management; Business environment*; *Governance; Risk analysis; Risk management strategy.*
- ✓ *Protect*: it is linked to the implementation of measures meant to protect the business processes and assets, regardless of their IT nature. The Categories are*: Access Control; Awareness and Training; Data Security; Information Protection Processes and Procedures; Maintenance; and Protective Technology*.

- ✓ *Detect*: it is linked to the definition and implementation of appropriate activities intended to identify IT security malicious events on time. The Categories are: *Anomalies and Events; Security Continuous Monitoring; and Detection Processes.*
- ✓ *Respond*: it is linked to the definition and implementation of appropriate activities intended in case of detection of a cybersecurity event trying to reduce the impact of a potential cyber-security event. The Categories are: *Planning; Communications; Analysis; Mitigation; and Improvements*.
- ✓ *Recover*: it is linked to the definition and implementation of activities intended to the management of plans and activities, ensuring the resilience of systems and facilities.
- ✓ Categories within this Function include: *Recovery Planning; Improvements; and Communications.*

The Profiles embody the selection made by an organization upon specific *Sub-categories*.

The implementation Tiers provide the context to which companies considers cyber-risk and processes to manage it. There are four evaluation levels: *Partial, Informed, Repeatable, Adaptive.* They drive from the softest to the hardest one.

Shifting to the *priority levels,* they are meant to support organizations in the preliminary identification of *Sub-categories* that must be implemented to reduce their risk levels. The use of a priority scale of three levels midst *Sub-categories* is suggested by the Framework itself.

The identification of priority levels follows two specific criteria:

- ✓ The ability to reduce cyber-risk;
- ✓ The simplicity of *Sub-category* implementation.

Combining these two criteria permits the definition of three different *priority levels*:

- ✓ *High Priority*: actions that are prioritized and must be implemented irrespective of their implementation complexity;

- ✓ *Medium Priority*: actions that are generally easily implementable;
- ✓ *Low Priority*: actions that are generally considered hard to be implemented (deriving from significant organizational and/or infrastructural changes).

Moving to *Maturity levels* it is possible to state that they enable the measurement of maturity of a security process, a specific technology implementation or an assessment regarding the amount of resources needed to implement a definite *Sub-category.*

Dealing with *Framework contextualization* means to create a contextualization of the Framework to select the *Function, Category and Subcategory* of the relevant *Framework Cores*, specifying the appropriate *priority and maturity levels* for the *implementation context (*mandatory, recommended, free*).*

The *GDPR contextualization* means to adapt to a *specific contextualization prototype* that captures the elements fundamental of the GDPR [6]. The fundamental elements would be:

- ✓ Roles and responsibilities;
- ✓ Registers of processing activities;
- ✓ Principles and accountability;
- ✓ Impact assessment on the protection of personal data;
- ✓ Information of the interested party;
- ✓ Consent of the interested party;
- ✓ Rights of the interested party;
- ✓ Transfers of personal data to third countries or international organizations:
- ✓ Management of incidents that are configured as personal data breaches.

It is also accompanied by an implementation guide, a reference sheet including: the context of application of the chosen prototype, further additional constraints on the selection of subcategories if needed, and the definition of priority levels. It is also important to assess that

several prototypes can be implemented at once in contextualization phase to capture different aspects (i.e. GDPR [6], NIS [18], CCS-CSC [12]).

The inclusion of the *Framework Core* with new categories and subcategories (DP- as for Data Protection) related to the GDPR *prototype of contextualization* clarifies which are the categories of controls fundamental for the compliance with GDPR [6] and, in detail, reports for each category the article of the European Regulation concerned, to obtain an adequate support tool for organizations, security managers and, Data Protection Officers that need strategies and processes aimed at protecting personal data and cybersecurity.

3 Ospedali Riuniti Case Study

Healthcare is one of the main industries concerned by the General Data Protection Regulation (GDPR) [6]. This is mainly due to the severity of the risk that a data breach and, in general, a cyber-attack, could potentially have on subjects. Therefore, our case study is strictly correlated to a specific category of data that is generally processed within hospitals and clinics, so healthcare facilities. Article 9 (Processing of special categories of personal data) suggest the adoption of advanced protecting measures, introducing further conditions and limitations when processing one or more specific categories of data (including genetic, biometric and, health data). It also indicates whether and how an organization should process personal data.

This latter constraint is strongly inspired by the principle according to which data itself does not constitute information: its processing does, increasing informative content of data itself [16]. Hence, the organization is allowed to process data only if consented explicitly by the subject and/or strictly necessary, as stated by Article 6 (Lawfulness of processing) of GDPR [6]. Hence, an organization is invited or even enforced, to maintain records of data processing activities in order to certify which and how personal data are processed.

3.1 Records of data processing activities

The record of data processing activities is a tool introduced by GDPR [6] and amply addressed by Article 30 (*Records of processing activities*). it defines when maintaining the record is mandatory and the minimum content. It also states that the record must be maintained in electronic format. The Data Controller is enforced to draft the record when at least one of the following criteria is verified [16]:

- ✓ societies or organizations with more than 250 workers;
- ✓ processing information that could result in a risk to the rights and the freedom of the subject;
- ✓ processing information in a non-occasionally way;
- ✓ processing sensitive categories of data (Article 9) or data about criminal convictions or crime (Article 10).

Due to those latter criteria and requirements healthcare facilities are always enforced to draft records of data processing activities.

The fields of the records processing activities required are the following [16]:

- ✓ the name and contact details of the data controller and, where applicable, the joint controller, the controller's representative and the data protection officer (DPO);
- ✓ the purposes of the processing;
- ✓ a description of the categories of data subjects and of the categories of personal data;
- ✓ the categories of recipients to whom the personal data have been or will be disclosed;
- ✓ where applicable, transfers of personal data to a third country or an international organization;
- ✓ where possible, the envisaged time limits for erasure of the different categories of data;
- ✓ where possible, a general description of the technical and organizational security measures.

The records of data processing activities have been introduced by GDPR [6] as one of the main tools to accomplish accountability. The Data Controller, through the records of data processing activities should demonstrate compliance with GDPR in a proportional way to the risk associated to protect personal data. The organizational and technological measures adopted by Data Controller should be chosen to guarantee the protection of information.

The aim of the current study is to accomplish the beforementioned complex requirements thanks to a new tool proposed in the next paragraph.

3.2 Design of a new tool

As suggested by the regulation, the records of processing activities should be draft in electronic format allowing the Data Controller to update and implement changes. A smart approach is to exploit the new tool to accomplish the requirements imposed by the Regulation. The records of processing activities should perform *risk assessment* because a risk-based approach is highly recommended by the Regulation. By computing a risk index, the record evolves from a *qualitative tool* that aims to demonstrate compliance to the Regulation, to a *quantitative* one that allows the controller to monitor and manage the risk, strengthening the compliance to GDPR.

3.3 Tool's Framework

Since we have defined the aim and the possible format of the new tool, the successive step should be the assessing of a suitable framework. The data processing activities should be organized in six linked sections as follows:

- ✓ ***Cover:*** it contains the name and contact details of the Data Controller and, where feasible, the joint controller, the controller's representative and the Data Protection Officer (DPO);

- ✓ ***Record of data processing activities***: it is composed by a specific module, one for each data processing activity. The module contains all the information required, as well as some descriptive fields;
- ✓ ***Asset & technological context***: it contains the list of softwares, and IT infrastructure involved in personal data processing within the organization;
- ✓ ***Technical and organizational measures***: it contains a list of all possible technical or organizational measures used to mitigate the risk. Each of them is described and assigned with a score that represents the associated risk reduction for confidentiality, integrity and availability of data;
- ✓ ***CIA triad:*** it contains the list of categories of data with a score about confidentiality, integrity and availability assigned to each category: category of data, data processing activity, automaticity of processing, numerosity of processing, frequency of processing, category of subjects, age of subjects, storage time and receivers;
- ✓ ***Info:*** it contains description and development of the *risk assessment*.

3.4 Principles of General Data Protection Regulation [GDPR]

After the introduction and widespread of GDPR [6], most companies, organizations and operators have found themselves astonished, at least at the beginning, with the Regulation. As far as it goes, it is significant to say that the GDPR should not be seen as an obstacle to innovation or yet another, rigid, bureaucratic fulfillment by the operators, but as an organizational evolution, aimed at greater protection of the freedoms of those concerned and involved. GDPR [6] tries to improve the current state of the art, especially in processing personal data. The principal aim of the EU is dual: firstly, the protection of rights and freedom of European citizens given by a new regulatory framework; secondly, as a direct correlated consequence of the first aim, the adaptability to the rapid development of the information society to personal data encouraging the innovation that brought inherently.

The first difference that must be highlighted is the one concerning the discrepancy betwixt personal data and data profiling, i.e. data processed through treatments since GDPR deals with the second one exposed as a form of automated treatment. Hence, data profiling requirements usually tend to evaluate or judge the interested party, dealing with data, as contain a major amount of information than personal data [16]. Therefore, every definition of personal data, or simplistic as data, has to be interpreted as personal information. It is important to assess that, the key element of GDPR is not personal data itself, but its processing through treatments.

An important aspect to consider is the compliance with GDPR [6], not so feasible to achieve because of its requirements, even though is not so strictly from a technical and judicial point of view. The Regulation statuses some guidelines and principles that the Data Controller (the physical or legal person who has the power to decide means and purposes of processing personal data) must follow and achieve. GDPR undoubtedly constituted an aggravation concerning the obligations borne by companies, professionals, and public bodies in the field of personal data protection and processing [19]. To identify the most suitable technical and organizational measures, it is necessary to list and cite GDPR basic concepts and definitions [16]:

- ✓ ***Privacy by design*****:** the Data Controller must put in place the finest data processing assessment method starting from the definition of data processing.
- ✓ ***Privacy by default*****:** the assessed data processing must provide, by default, only strictly sufficient personal data and they must be conserved for the shortest time available, in line with the purpose of the processing.
- ✓ ***Accountability***: the Data Controller must select and adopt the best security measures based on the state of the art. Also, it must update these measures and be able to demonstrate continuously the compliance of its organizational model assessment.
- ✓ ***Risk-based approach*****:** the Data Controller must preventively evaluate every kind of risk associated with data processing, pointed out the ones concerning data-breach and non-authorized diffusion of themselves.

The reference articles of GDPR related to the risks of processing personal data *Art. 32 (Security of processing*) and *Art. 25 (Data protection by design and by default*) are shortly listed below [16]:

- ✓ "*Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organizational measures, such as pseudonymization, which are designed to implement data-protection principles, such as data minimization* […]" (Art. 25.1)

- ✓ "*In assessing the appropriate level of security account shall be taken in particular of the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to personal data transmitted, stored or otherwise processed*." (Art. 32.2)

3.5 The Treatment Register of Ospedali Riuniti

Is in this very specific scenario that our case study, regarding "Azienda Ospedaliero - Universitaria Ospedali Riuniti of Torrette, Ancona" situates.

On 15 May 2018 the Data Controller, represented by the DPO in our specific case, of Azienda Ospedaliero-Universitaria Ospedali Riuniti of Ancona drafted its first and actual "Treatment Register" including 201 data processing activities. The Data Controller has been helped by an engineer working for the SIA's ward (Corporate Information System) and an external expert to give a validation to the document proposed.

The initial aim of the work was to analyze the actual document, the “Treatment Register” and all the data processing activities involved to solve inadequacy or discrepancies that could have arose.

The “Treatment Register” is formed by an Excel document, comprised of 9 Sections or sheets listed in order of appearance:

- ✓ Name of the data controller, legal representative, personal data protection officer, method of conservation of the actual register;
- ✓ RT Administrative area containing all the processes and treatments pertinent to that specific area;
- ✓ RT Healthcare Area;
- ✓ RT Services Area;
- ✓ RT Hospital Medical Management;
- ✓ Sections containing categories of data involved, technical and organizational security measures and categories of subjects;
- ✓ Healthcare organization chart with a description of the assignments present of each macro-area
- ✓ Section containing a list of the entire amount of processes occurring in each ward;
- ✓ Administrative organization chart;

The sections 2 to 4 are the ones concerning all the data processing activities involved and their specific partition is described as follows:

- ✓ Classification/Taxonomy:
    i. Macro-area considered;
    ii. Area as a subsection;
    iii. Organizational structure of the entire hospital;
    iv. Responsible of the treatment;

- ✓ Processes;
- ✓ Treatments;
- ✓ Title of data processing activity;
- ✓ Legal basis and contextualization;
- ✓ Aims of the treatment;
- ✓ Storage time;
- ✓ Nature of data (sensitive or not, based on Art. 9 of GDPR [6]);
- ✓ Categories of data (listed in detail also in Section 7);
- ✓ Categories of subjects involved;
- ✓ Information and consent on data processing;
- ✓ Categories of recipients;
- ✓ Technical and organizational measures adopted;
- ✓ Data processing activity co-owner;
- ✓ External responsible (when feasible);
- ✓ Data transfer (UE, not UE).

At first glance, the organization of the "Treatment Register" seemed to have several critical discrepancies resulting in a lack of mandatory important information. Its division is merely based on organizational or administrative categories, without any defined and specific hierarchy. Besides, there seemed to be no *quantitative assessment*, but only a defective *qualitative assessment* without numerical approaches or standardizations. Indeed, the list of treatments and processes also showed an excessive aggregation of data, ending in a possible loss of fundamental information necessary for the purpose of the risk analysis.

After a discussion on the actual register has been carried out, all the problems and discrepancies listed and internalized, it has been decided according to the DPO, the external coadjutant expert, and SIA's representative to try a new tool for the assessment of cyber-risk in healthcare.

The underlying idea and all the potentialities and functionalities of this new tool have been well illustrated and explained to the DPO, to the external coadjutant expert, and to the SIA's representative engineer. All the parties seemed to agree with the idea of changing the actual "Treatment Register" trying to overcome the problems that have arose.

The aim of the work was to achieve a *quantitative assessment* starting from a defective *qualitative assessment,* trying to calculate numeric values associated with each component (treatments or processes in our case) that result after the risk evaluation.

3.6 Critical Problems Encountered

The principal aim was a review of the actual improper taxonomy of the register. We tried to map and subdivide the "Treatment Register" according to the purpose of processes instead of the treatment itself (several misunderstanding about the distinction between processes and treatments), according to a hierarchy based on the area (technical, administrative, healthcare, services) considering the technological context. Our final aim was the evaluation of a risk assessment, so the first step included the evaluation of the purpose of the treatments linked to the technological asset encountered. Given the critical fragmentation of the "Treatment Register", we realized that our attempt would not have been possible at all. The document lacks necessary data about the technological asset of the business considered not considering information and communication technologies development level at all. The original hierarchy map, obtained with a software (Xmind 10.2.1), is shown below:

Figure 2: Taxonomy of Ospedali Riuniti divided in Macro-areas.

The original mapping was useless for the purposes of the risk assessment, treatments and processes are not well differentiated and, in several cases, seemed to be conflicting, the external responsible was no mention at all even if has responsibilities over several processes. In addition, the description of come treatments often does not correspond to the purposes or to the processes involved, lacking punctuality and necessary information.

At this specific point, we tried to exploit the new assessment tool inserting the entire amount of data from the previous model assessment to the new one. The two methodologies seemed to be consistent and coherent, but several fields have discrepancies, especially the critical ones: context and adopted technical and organizational measures are missing, co-owner has never been established, the likelihood In case of automatic or manual process is not attained.

3.7 Insight of the "Treatment Register"

In this paragraph a single record of the "Treatment Register" is considered and shown (*Figure 3*), in order to consider and evaluate all the problems discussed in the previous section, to present the

actual information inserted in the “Treatment Register” within their validation and, last but not least, to evaluate how the tool behaves when computing the risk index assessment.

Data processing activity	**Taxonomy**	RT Administrative AREA			
		Functional economic-financial and administrative flows			
		Company Information System			
	Title of processing activity	Patient Registry			
	Processing activity ID	42			
	Purpose	Ensure the synchronization of the company registry with the regional one University Hospital of Ospedali Riuniti (SIA)			
	Data controller	Azienda Ospedaliero Universitaria Ospedali Riuniti di Ancona			
	Co-owner	None			
	External responsible	None			
	Description	Management and updating of patient registry (MPI) with synchronization with regional registry (ARCA)			
	Data processing activities	**C**	**I**	**A**	
		1	1	1	Recording
		1	1	1	Storage
		1	1	1	Consultation
		1,1	**1**	**1**	Communication
	Frequency	**1**	**1**	**1**	Standard
	Numerosity	**1**	**1,1**	**1,1**	High
Law	**Lawfulness**	Subject’s explicit consent			
	Information model	Model defined by the project			
	Consent model	Model defined by the project			
D.p.a. object	**Category of personal data**	2	2	2	ID data
		8	6	8	Health data
		10	**8**	**8**	Genetic data
	Subjects	**1,1**	**1,1**	**1,1**	Patients
	Age of subjects	**1,2**	**1,2**	**1,2**	Adults and/or children
	Storage time	**1,0**			Reference law
Recipients	**Recipients**	**0,9**	**0,9**	**0,9**	Internal
		1,2	**1,0**	**1,0**	SO, Company Information System
	List of internal recipients	SO, Company Information System			
	Extra-UE recipients	**1,0**	**1,0**	**1,0**	None
	List of extra-UE recipients	Void			
Te	**Software/application/MD**	0,5	0,5	0,5	Void
		0,5	**0,5**	**0,5**	Business IT systems

	IT infrastructure	**0,3**	**0,3**	**0,3**	Informatic archive
	Likelihood (automatic process)	**0,4**	**0,4**	**0,4**	
	Paper archive	**0,7**	**0,7**	**0,7**	Protected archive
	Likelihood (manual process)	**0,7**	**0,7**	**0,7**	
Privacy chain	**Person in charge**	Stefano Occhiodori			
	Area of person in charge	Company Information's System			
	External person in charge	None			
	Date of contract	Void			
	Expiring date of contract	Void			
	Contract	Void			
	Sub-responsible	None			
Risk assessment	**Impact of category of data**	**10**	**8**	**8**	
	Impact of amplification factors	**1,6**	**1,3**	**1,3**	
	IMPACT	**16**	**8,2**	**8,2**	
	LIKELIHOOD	**0,57**	**0,57**	**0,57**	
	RISK INDEX RI = L*I	**9,1**	**4,7**	**4,7**	
		Very high	Medium	Medium	
	Technical and organizational measures	**1,0**	**1,0**	**1,0**	None
	RISK INDEX considering adopted measures	**9,1**	**4,7**	**4,7**	
		Very high	Medium	Medium	
Notes	**Date of assessment**	19/11/2019			
	Insights to do	Poor information on technological context; Very high risk: adopting technical and organizational measures is extremely necessary			

Figure 3: data process activity of one record in the "Treatment Register"

It is important to assert that when evaluating the *risk index assessment,* the values considered are the one that possess higher value for each category considered, hence, maximum risk within the pertinent category since lesser value would not influence the risk index as they do not represent the maximum risk index obtainable regarding the specific data activity.

3.8 Risk analysis performed by the new tool

Risk indicators are usually based on the probability of occurrence of an event and the effects it could have on the person concerned. the algorithm proposed below envisages using the information entered in the processing register as input in order to systematically calculate a risk index for the data subject. the algorithm is based on the following simple calculation (already seen in the processing register):

$$\boldsymbol{RI = L * I}$$

Where:

- ✓ $\boldsymbol{RI = risk\ index}$
- ✓ $\boldsymbol{L = likelihood\ of\ occurrence\ of\ an\ adverse\ event}$
- ✓ $\boldsymbol{I = impact, as\ severity\ of\ consequences}$

The computation of the risk index on the purpose of the risk assessment evaluation could not be considered statistically consistent because of the constraints listed and explained in Paragraph 3.3.

In particular, the necessary data missing in the tool are:

1. Software information and maturity index:

    i. The quality of application software development processes *(Privacy by design and Privacy by default);*

    ii. The quality of the processes related to the use of the software.

2. Information about technological chain and maturity index of the processes:

iii. The quality of the technologies and architectures underlying the software system (IT system);

iv. The quality of the processes with which the technologies are managed and maintained (IT system management).

4 New Assessment Model

At the beginning of any risk assessment approach, it is essential to define the scope and subject of the assessment. These, in fact, are used to determine what type of risk we are evaluating. The impact, or the set of consequences that occur upon the occurrence of an event, is also closely linked to the scope of the risk [20]. Numerous approaches of risk assessment have been proposed in literature which can help identify risk, assess the risk appropriately and support in the *Risk Management* [21]. Many of these approaches have the same goals as this new assessment model, but they are often useless complicated and difficult to interpret. As far as concerned pre-existing assessment approaches fail to capture the essential elements necessary for the correct assessment of this type of risk.

The newly proposed approach is focused on several key elements, which would be explained as long as the thesis goes. The evaluation of an $complexity\ index$ of the considered IT infrastructure along with the $maturity\ index$ (obtained through calculation starting from one or more existing reference frameworks considered in *Section 2*) are two of the essential key elements necessary to correctly assess the risk supporting *Risk Management*. Hence, it is possible to understand that the resulting risk will not be complementary to ventures' maturity index (in a way that a greater maturity index would correspond to lower overall risk, so an inverse proportionality) but instead would show a direct proportionality between the variance existing between the complexity of the context analyzed and the organizations' $maturity\ index$. Therefore, having the same risk index would mean that firms' maturity index must increase as the complexity of the infrastructure analyzed increases, as a result of direct proportionality. Organizational attractiveness represents an additional important feature that sometimes-pre-existing risk model assessments fail to interpret or to evaluate at all, although has become an increasingly important concept in literature [22]. Along with the other parameters involved, the organizational attractiveness would influence the likelihood of occurrence of an adverse event within the organization considered. The likelihood of occurrence of an adverse event would represent the aim of our model assessment methodology, thanks to standardized quantitative estimations, as cyber-security threats have exposed individuals

and organizations to a host of new risks resulting from attacks through digital interfaces [23]. Therefore, the proposed approach has the goal to effectively partially or entirely overcome the assessment methodologist problems previously pointed out, combining the evaluation of the maturity index of the processes inherent to firms along with the evaluation of the complexity index of the desired firms with chosen infrastructure concluding with the attractiveness of the organization.

4.1 Complexity Index

The Center for Internet Security (CIS) is one of the forerunners in developing guidelines for protecting people, organizations, and governments from cyber threats in our continually evolving digital landscape [11]. CIS critical security controls function based on risk assessment best practices and contain guidelines to provide the proper maintenance, monitoring, and analysis required to secure an organization. Through the development of CIS Controls V7.1 and the Implementation Groups (IGs) [11] it would be possible to define self-assessed categories for organizations based on relevant cybersecurity attributes. Its aim is to provide a simple and accessible way to help organizations of different classes focus their security resources, and still leverage the value of the CIS Controls program, community, and complementary tools and working aids. The CIS Controls Implementation Groups (IG) fall into three categories based on appropriate cybersecurity attributes. IGs have their respective subset of controls, estimated to be executed reasonably and affordably. Each IG is more complex than its predecessor and scales based on an organization's size, type, and function. All these IGs require the analysis of audit logs to prove compliance and secure configurations of hardware. It is easily comprehensible that every organization should identify itself, mainly according to its dimension, in one of three IGs through a general self-assessment and, consequently, needs to implement different subsets of controls. Accordingly to the IGs to which every venture belongs also data sensitivity and criticality of services offered by the organization would be influenced, by the expected level of technical expertise exhibited by the staff, and by the resources available and dedicated toward cybersecurity activities. In fact, a differentiation like the

one proposed by CIS, introduces the key concept that the maturity index of a specific firm would be strictly corelated to the complexity index of its entire infrastructure.

Our approach evaluates and tries to calculate the intrinsic complexity of the IT infrastructure chosen in an objective and punctual way, avoiding subjectivity and possible ambiguities. The entire subsets of controls utilized to assess the complexity index of ventures might be classified into 5 aspects: *Networks and Infrastructure, IP network technologies, Applications, Services and IT department*. Starting from the detail that the existing literature fails to determine correctly and objectively a proper complexity assessment framework, the subsets controls have been selected considering the possible critical points regarding hardware, software, networks, and facilities. Number of components (physical and software systems), characteristics of the components (level of obsolescence, certified contexts of use), interconnections between components, number and characteristics of services, and entropy in IT System management have been so far considered. What would be the desired outcome look like? It would be calculated through a weighted average of the complexity scores for each of the five above-mentioned categories and aspects, as an index ranging from 0 to 10, describing the intrinsic complexity of the IT infrastructure chosen.

4.2 Maturity Index

When evaluating the maturity index, a fundamental correlation is that the measure of the level of adherence of the firms to the subsets of controls, specified by the chosen reference framework, corresponds to maturity index itself. The first step a venture has to follow is to perform a prior assessment method grounded on the subsections of controls regarding one or more pre-existing frameworks (i.e. CIS, ENISA, ISO, FNCS). The choice relies on the fact that after having chosen a specific framework with the correlated subsets of controls inherent to the framework itself, the area of application of the model is easily determined and ready to be examined. For example, exploiting the CIS controls [6], firms would determine their compliance and adherence under cybersecurity terms, on the other hand

ENISA controls [10] would focus compliance and adherence of a venture on data protection. Shifting to FNCS controls [7], they would help an organization evaluating its compliance in both cybersecurity terms and data protection appliances, comprising them both.

The idea underling the usage of pre-existing frameworks to determine maturity index relies substantially on the staple that pre-existing frameworks are usually considered as a set of well addressed and solid best practices in a way they might provide a complete mapping of the area of interest, typically sided and affiliated by official correlations among different standards. The complete subsets of controls of the selected framework could be interpret as complementary controls to the list of all the possible adverse threats. Parallel to the declaration, the evaluation of the maturity index on every subsets of control implementation leads to an almost complete analysis of the vulnerabilities of the infrastructure considered. In fact, starting from the former quote, the evaluation of the possible adverse threats and the list of the existent vulnerabilities through the complementary subsets of controls of the reference framework would spring in an investigation based on the actual posture of the society, so the model assessment would consider not only a list of past experienced threats bus also emerging new ones. Furthermore, as we already be aware of that reference pre-existing frameworks are constantly up to date, their exploitation leads to the creation of a dynamic model; specifically societies might be able to update their actual posture by simply repeating the maturity assessment every time a new version of the reference framework is released. The selected assessment method to calculate index maturity is based on the subsets of controls offered and clarified by CIS Controls V7.1 Implementation Groups [7] as our preferred reference framework. So, accordingly to CIS Controls V7.1 [7] we could define self-assessed categories for the desired organizations based on relevant cybersecurity attributes, in particular our assessment method would focus on attacks against systems and networks. The outcome of this assessment method is a maturity index, ranging from 0 to 10.

4.3 Attractivity of ventures

Let's switch to the last parameter we would like to consider in this model assessment: *attractivity of ventures*. Attractiveness basically means to cause interest or pleasure and to pull someone towards you by the qualities you have, especially positive and admirable ones [24]. The attractivity of ventures could be referred to the attractiveness that a particular company possesses from possible attacking cyber criminals especially when there is an attractive pay-off. It usually depends on several and disparate factors such as: the kind of business taking in consideration, the type of processed data (sensitive, protected, processed, etcetera), the organization purpose and scope in market business, and so on. It is easy to understand that different organization have different attractiveness, and, consequently, are exposed to dissimilar levels of risk and, attackers. Another important aspect is that in a precise defined time interval, a lower appealing organization (defined by a potential casual attacker) will be subjected to a reduced amount of attacks if compared to a higher appealing organization; moreover, each attack would be comprise of fewer breaching attempts in case of a lower appealing organization, than the ones showed by a higher appealing organization; they would receive major and more structured refined amount of attacks. Thus, the attractivity of venture is strictly correlated to the concept maturity of the adverse attacks and of the criminal attackers: a more appealing organization (in terms of economic benefits but also dissimilar aspects) would have a greater probability of being attacked.

The proposed approach has examined its attractivity of firms exploiting data proposed by CLUSIT in 2020 regarding IT security in our country [11]. In particular, different types of ventures have been considered along with the possible adverse attack received and suffered by the proposed categories in 2019, so basically it has been considered the distribution of the victims of attacks divided by the kind of organization. Starting from this specific point, the attractiveness of a sort of organization was estimated and calculated, classifying it in *Very Low, Low, Medium, High, Very High*, considering how much the amount of attacks suffered by that specific kind of organization differed from the average of the attacks suffered by each typology during the year considered.

4.4 Model Components

The likelihood and occurrence of a possible adverse event is influenced by the maturity of the attack (dependent, in turn, on attractiveness), by the maturity index and complexity index of the organization.

4.4.1 Ventures' Posture

The maturity and complexity indexes of a specific venture give an actual description of its posture affecting nevertheless the likelihood and occurrence of a successful adverse attack.

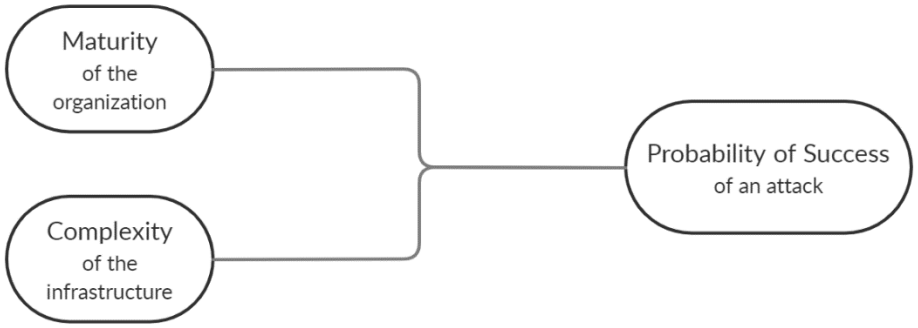

Figure 4: relations holding between maturity of the organization, complexity of the infrastructure and probability of success of an attack

Let's give an explanation about the parameters we have just seen before:

- ✓ *Complexity of the Infrastructure - Probability of success of an attack:* It is easy to understand that the complexity of the infrastructure and probability of success of an attack are directly correlated: hence, an increase in the complexity of the infrastructure would lead to an increase itself in the probability of success of an adverse event and, vice versa, a decrease in complexity of the desired infrastructure would lead to a successive decrease in the likelihood of an adverse event being successful.

- ✓ *Maturity of organization - Probability of success of an adverse attack*: the maturity of an organization and probability of success of an adverse attack are, in this model assessment, inversely correlated: as a matter of fact, an increase in maturity of an organization would prime to a lower probability of success of an adverse attack and, consequently, to a decrease in risk, and, in the opposite way, a lower maturity of an organization corresponds to a higher probability of success an adverse attack.

4.2.2 Context of the Organization

The context of the organization determines its attractiveness which therefore influences the number of attacks and the type of attackers.

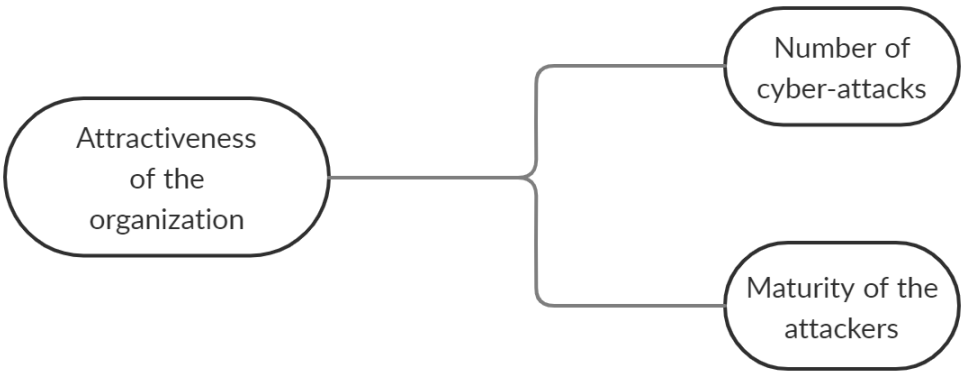

Figure 5: Relations between attractiveness of the organization, number of cyber- attacks and maturity of the attackers.

Let's give an explanation about the parameters we have just seen before correlating them in couple as the model proposed:

- ✓ *Attractiveness of an organization - Number of adverse attacks*: the attractiveness of an organization is directly correlated to the number of possible adverse attacks as it could be

seen before. Higher attractiveness of an organization leads to more recurrent adverse attacks and, conversely, a lower attractiveness of an organization will be subjected to fewer adverse attacks.

- ✓ *Attractiveness of an organization – Maturity of cyber attackers:* the attractiveness of an organization is also directly correlated maturity of cyber attackers. Higher attractiveness of an organization tends to attract experienced and well-prepared attackers launching more structured and targeted adverse attacks and, conversely, a lower attractiveness of an organization would tend to attract less experienced cyber criminals preparing simpler attacks.

4.2.3 Probability of occurrence of a cyber-attack

The number of cyber-attacks, along with the likelihood of successfulness of a cyber-attack, affects the likelihood of occurrence of a cyber-attack. The likelihood of successfulness of a cyber-attack, however, is also affected by the maturity of cyber-attackers.

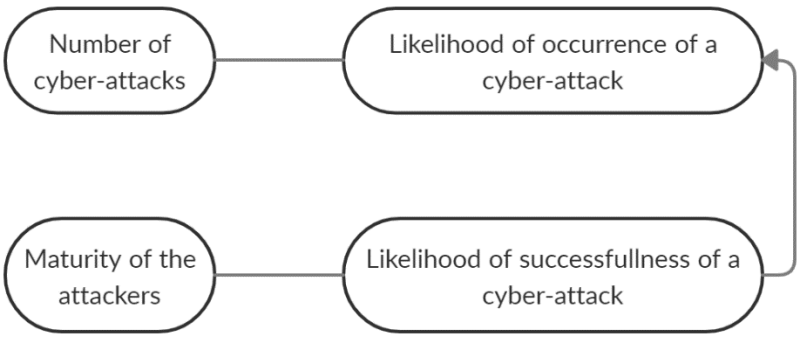

Figure 6. Factors *affecting* the likelihood of occurrence of a cyber adverse event.

Let's give an explanation about the parameters we have just seen before correlating them in couple as the model proposed:

- ✓ *Maturity of Attackers – Likelihood of Success of a cyber-attack:* the maturity of attackers is directly correlated to likelihood of success of an cyber-attack, hence, well experienced cyber criminals would perform more structured and targeted cyber-attacks and, therefore, the likelihood of success of a cyber-attack would be greater lowering the likelihood of successfulness of simple and non-targeted cyber-attacks. An assumption conceivable is that an attack is precisely comprised of a series of preparatory attacks making possible to acquire even more precise and complete information before conducting the possible final attack. Consequently, the higher the maturity of the attackers, the more preparatory attacks will be performed having a huger likelihood of success of the sequence of adverse attacks. If the attacked organization notices the sequence of attacks in time and changes its posture, it can suddenly make the likelihood of success of the sequence of adverse attacks tend to zero.

- ✓ *Number of adverse attacks - Likelihood of occurrence of a cyber-attack:* the number of adverse attacks suffered by the organization would inherently and directly affect the likelihood of occurrence of suffering a successive successful attack. It is easily understandable that as the number of adverse attacks grows, the chances of having at least one successful adverse attack would rise up and, on the other hand, it is understandable that a venture subjected to fewer number of adverse attack attempts has a lower likelihood of occurrence of suffering a successful attack.

- ✓ *Likelihood of successfulness of an adverse attack - Likelihood of occurrence of a cyber-attack:* The likelihood of occurrence of an adverse attack is directly related to the likelihood of successfulness of a single adverse attack. If the cyber-attack is well-structured and has

potentially a high likelihood of success, the likelihood of occurrence of a cyber-attack is greater. Contrariwise, attacks with a lower likelihood of successfulness would result in a lesser likelihood of occurrence of a cyber-attack.

4.3 Structure of the model

The final structure of the model is exposed in the figure below comprised of all its relationships. The maturity and complexity of an organization together influence the probability of success of an attack. However, the likelihood of successfulness of an attack is also directly influenced by the maturity of the attackers, in turn linked to the attractiveness of the organization. Attractiveness also influences the number of attacks an organization will likely face over a given determinate period. The number of attacks, along with the probability of success of an attack, affect the likelihood of occurrence of an adverse event.

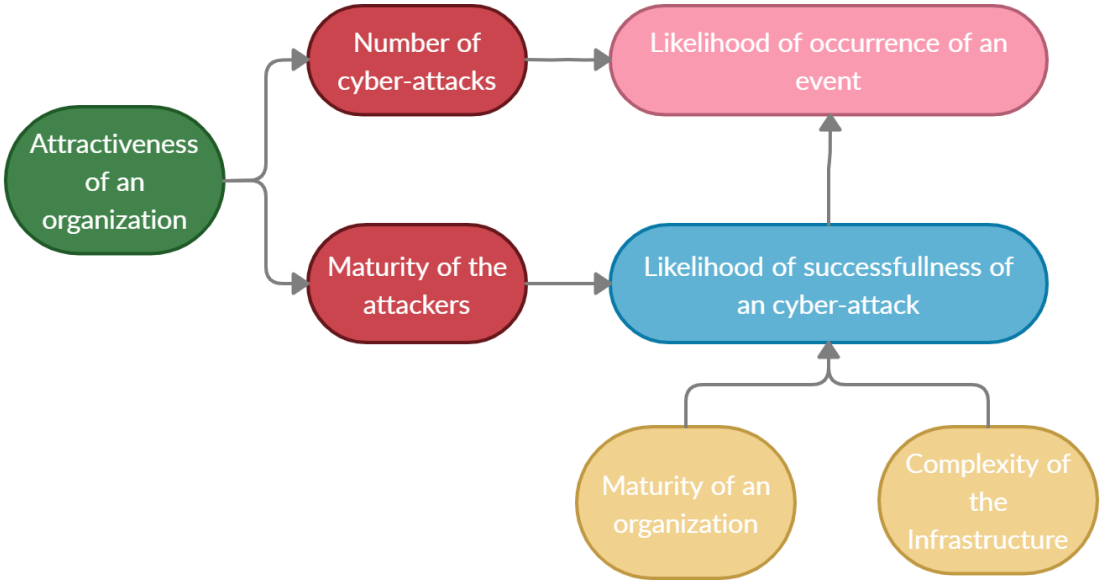

Figure 7. Summary of the relationships between the parameters involved in the model.

Considering also the impact of the event and therefore the risk we get the pattern depicted in *Figure 8*. It can be seen that an increase in the maturity of an organization leads to a decrease in the probability of success of an attack which therefore also decreases the probability of an adverse event happening and, consequently, the risk to which the organization is exposed

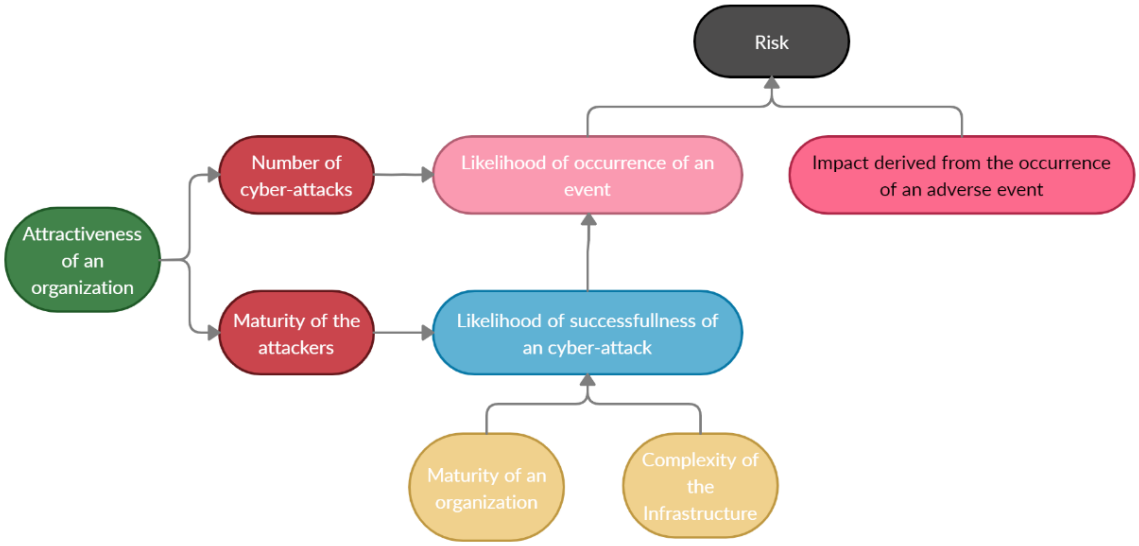

Figure 8. Structure of the model.

4.4 Assessing the model

In order to estimate in a quantitatively way the likelihood of occurrence of a cyber-attack, the goal is to build and to template a specific model in which, taking as input the complexity and the maturity indexes, will give as output the likelihood of successfulness of a single cyber-attack. The beforementioned likelihood of successfulness of a single cyber-attack would combine with the attractiveness of an organization (giving the number of specific injuries), thus might be used to estimate the likelihood of having an adverse event in a specified period. The proposal of this subsequent risk assessment framework might model the maturity of the stated infrastructure or venture and the relation betwixt the probability of occurrence of a malicious event (data breaches and/or attacks) through a logistic function. There are two important reasons behind the choose of

using the logistic function. Firstly, the logistic seems to describe the relation between the maturity of an organization and probability of occurrence of an attack in a precise realistic way, giving the model a fundamental assessment. As we already stated, the improvement in maturity organizations would give rise to a diminution in the probability of success in case of occurrence of a malicious attack, even if the correlation is not linear dependent. A hypothesis could be that the probability of success in case of occurrence of a malicious attack is more likely to be small when compared to levels of maturity far from the middle point, so for extremely high or low values of maturity organization. Secondly, we could exploit the generalized logistic function, also known as Richards' curve, [14] originally developed for growth modelling, as an extension of the logistic or sigmoid functions to better adjust the shape of the curve, in a more flexible way pretending to shape our function with the complexity index.

A logistic function (1) describes a sigmoid curve which increases exponentially for small values of x, and approaches a constant value asymptotically as x increases, and has been used extensively

$$\boldsymbol{F(x)} = \left(\frac{\boldsymbol{K}}{\boldsymbol{1+e^{-B(x-m)}}}\right) \quad (1)$$

Let's introduce and explain the parameters populating the function:

- ✓ *K= saturation level*, corresponding to the upper horizontal asymptote that limits and determine the curve's maximum value;
- ✓ *m = midpoint*, i.e. value of x corresponding to half the saturation level of the curve;
- ✓ B = growth rate, i.e. steepness of the curve.

Observing the trend of the logistic function (an example would be reported in *Figure 9*) three different stages could be identified: the initial growth, the exponential growth, and the deceleration.

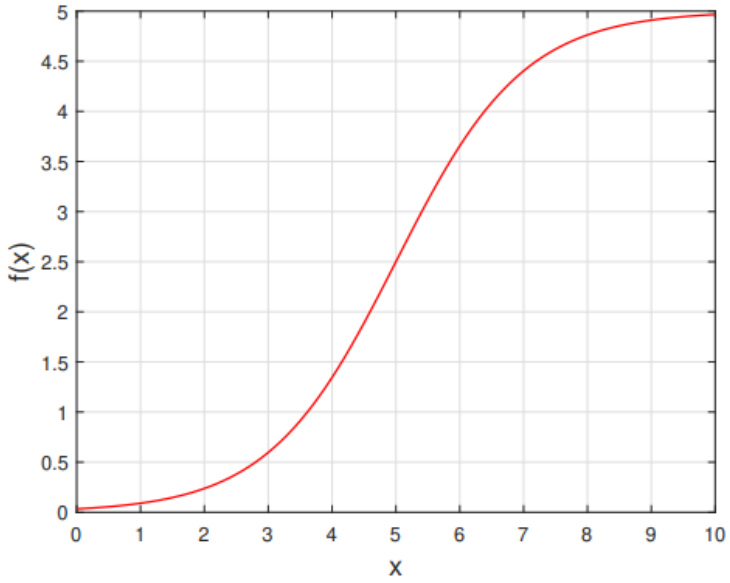

Figure 9: Logistic function with K = 5, m = 5, B = 1.

The subsequent plot (*Figure 10*) shows clearly how B, i.e. growth rate, affects the shape of curve showing the trend for different values of B, both positive and negative ones.

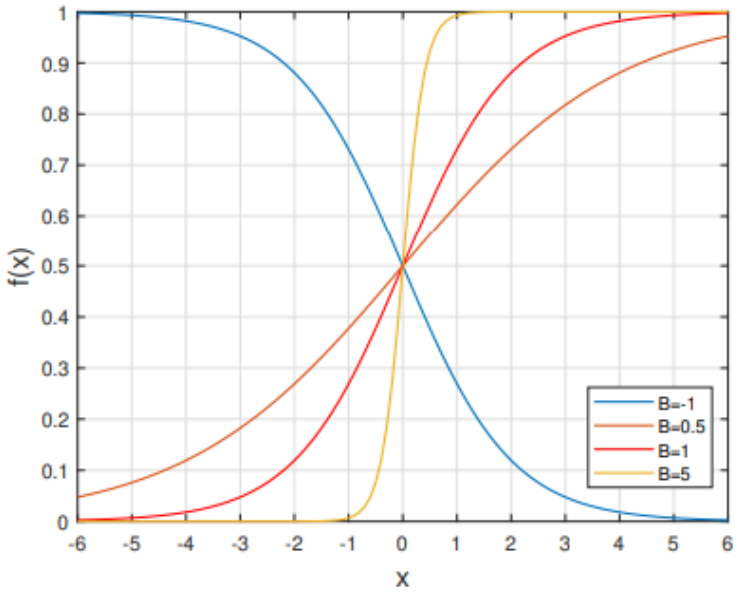

Figure 10: Effects of different growth rates B; (K = 1, m = 0)

For *K = 1, m = 0, B = 1*, the function become the "standard logistic function" and has is described by the following equation:

$$\boldsymbol{f(x) = \frac{1}{1+e^{-x}}} \qquad (2)$$

In practice, due to the nature of the exponential function e^{-x} , it is often sufficient to compute the standard logistic function for x over a small range of real numbers, such as a range contained in $[-6, \ +6]$, as it quickly converges very close to its saturation values of 0 and 1.

The beforementioned explained *standard logistic function* is used in the logistic regression to depict two variables having the sigmoid relationship, also known as "S"- shaped. Logistic regression (or logit regression) is a statistical fitting model that in its primary form tries to model a binary dependent variable exploiting the standard logistic function, although many more complex extensions exist. Therefore, it can be used to model the probability that a considered event occurs for a set of observations and the probability that the event does not occur. The logistic function can

be generalized introducing a lower asymptote different from 0 and a non-symmetric shape. The equation of the generalized logistic curve (or Richards' curve [12]) is the following:

$$f(x) = A + \frac{K-A}{\left(1+Q\times e^{-B(x-x_0)}\right)^{1/v}} \quad (3)$$

where:

- ✓ A = lower asymptote, i.e. the lower horizontal asymptote that limits and define the curve's minimum value;
- ✓ Q = variable related to f(0) that influences the inflection point;
- ✓ x_0 = punto in cui $f(x_0) = A + \frac{K-A}{(1+Q)^{1/v}}$;
- ✓ v > 0 determines the asymmetry of the curve.

The subsequent *Figure 11* shows how the variation v determines and changes the curve.

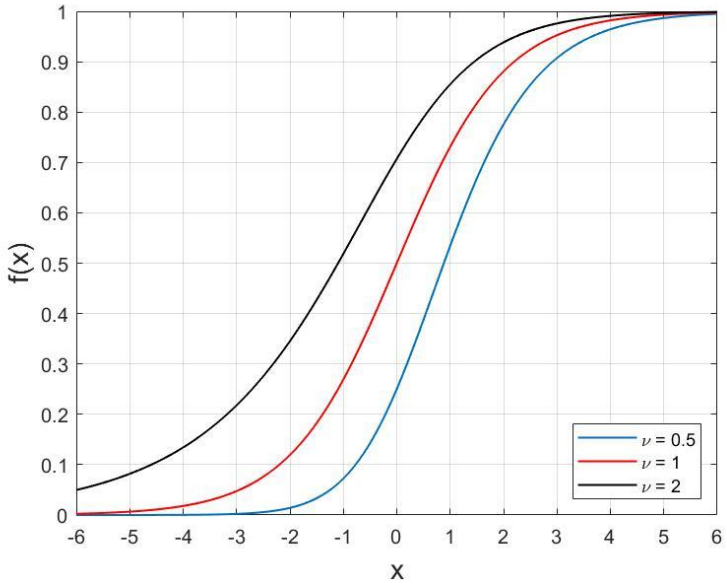

Figure 11: How v determines the asymmetry of the curve $(K = 1, A = 0, B = 1, m = 0)$.

In the assessment model proposed, we have fixed Q = 1 and v = 1: in this way, x0 corresponds to the point at which the curve is at its midpoint [13] and is the point of maximum slope of the curve. The final equation and subsequent reference model utilize (4):

$$f(x) = A + \frac{K-A}{\left(1+e^{-B(x-x_0)}\right)} \qquad (4)$$

The following *Figure 12* displays the behavior of a generalized logistic function.

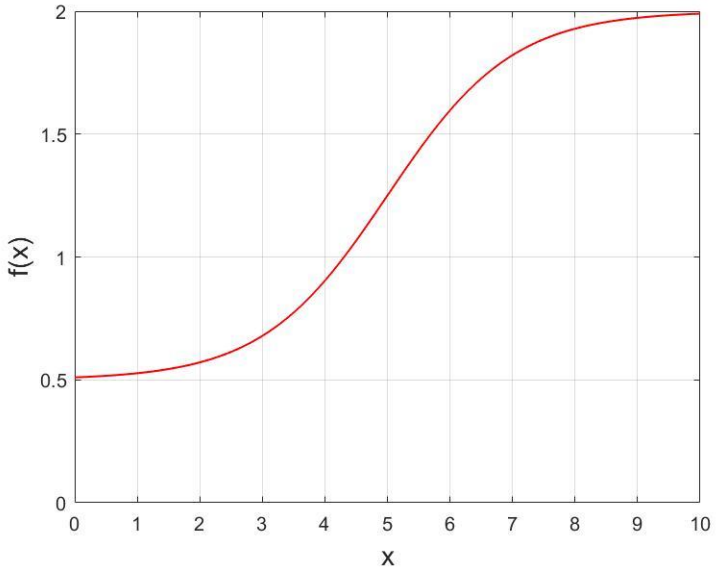

Figure 12: Generalized logistic function with $A = 0.5, K = 2, B = 1, x_0 = 5$.

Since our aim was to obtain a sigmoid curve statistically representative the correlation between the maturity of an organization with the probability of occurrence of a significative adverse event, we

elected the values of the parameter grabbed in consideration in a specific way in order to have a specific and well determined logistic function that was able to represent our data.

As we already seen before, the maturity of an organization and the likelihood of occurrence of a significative adverse event are inversely dependent such that an even slight increase of the maturity of an organization or venture would be correlated and coupled to a lower likelihood of occurrence of a significative adverse event, and off course the vice versa case clearly stands), so in this detailed case a negative growth rate value (we elite $B = -1$), but it was chosen a negative value as an example to obtain the preferred shape) was utilized in the logistic function. Our final aim would be to conduct a parametric analysis starting from real and specific available data in order to be able to elite the value of B perhaps in a heuristic mode finally proposing a binding methodology.

One assumption that has been pointed out it's that the probability of occurrence of a significative adverse event would never range a value equal to 0 nor 1. The beforementioned statement holds specifically because even assuming the worst-case scenario (the probability of occurrence of a significative adverse event would be close to 1 or the maturity of ventures would be the same In different circumstances), since the shaped logistic function would deal with real events concerning real people subjected to errors or deviances, it could not be avoided the fact that at the end nothing happened at all (i.e. the attack couldn't reach the infrastructure selected by the attacker/s), and reasoning the opposite way, also even if the maturity of ventures has reached its maximum value there is always a residual probability of a so-called black swan event, i.e. something unpredictable happening.

In order to better represent the model and the conditions we would shape and analyze, we decided to neglect the edges value, so 0 and 1, setting the maximum and the minimum quantity this way: 0.95 for $x = 0$ and at 0.05 for $x = 10$ respectively. Therefore, as we already stated that the measures of asymptotes K and A depend merely on the value of $x_{0,}$ it would be effortless to obtain them, starting and solving the following system of equations:

$$\begin{cases} f(0) = A + \frac{K-A}{1+e^{-x_0}} = 0.95 \\ f(10) = A + \frac{K-A}{1+e^{(10-x_0)}} = 0.05 \end{cases} \quad (5)$$

Terminating, the parameter x_0 has been set equal to the complexity index, previously derived from the complexity evaluation. Therefore, x_0 changes dynamically depending on the infrastructure considered. In this way, an increase in the complexity index is correlated to the urge of having a high value of the maturity index in order to contain the likelihood of successfulness of cyber-attack.

The subsequent resulting curves, for different values of x_0 could be seen *Figure 13*. Once a maturity index value has been settled, the increase in the complexity index, and the successive increase of x_0, is followed by an increase in the likelihood of successfulness of an adverse event.

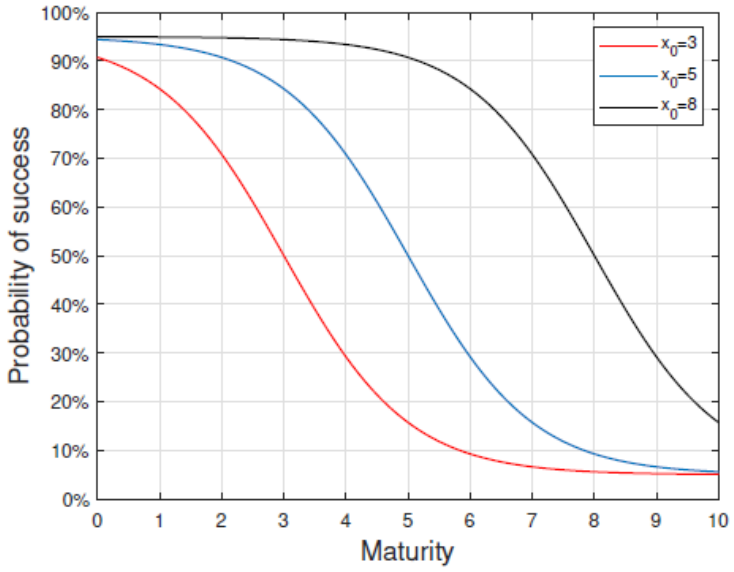

Figure 13: Effects of the change of x_0 in the proposed generalized logistic function.

Proceeding step by step, it is possible to comprehend that after evaluating the complexity index of the infrastructure considered, it is possible to define the shape of the probability curve. Consequently, considering the result of the maturity assessment, it is possible to easily obtain the probability of success (P) of a damaging event for the specific infrastructure considered. An example is shown in *Figure 14.*

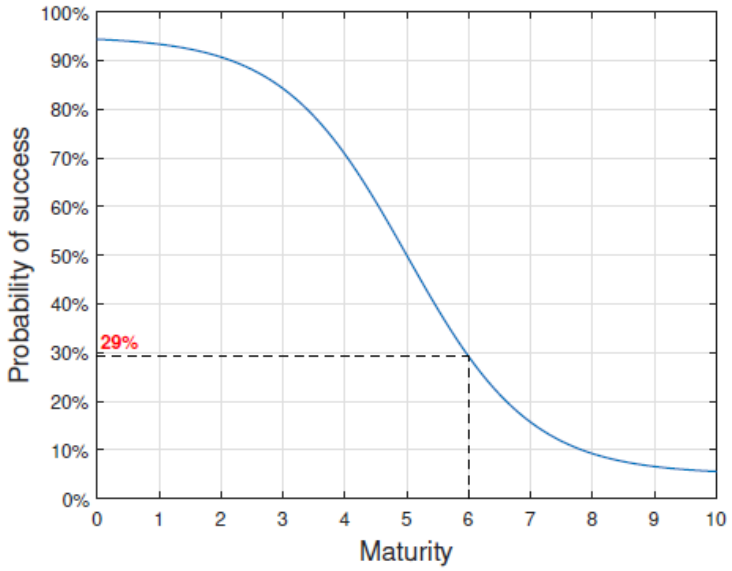

Figure 14: Example of how to determine the probability of success: $x_0 = complexity = 5; maturity = 6.$

Another parameter proposed and explained before, the attractiveness of the organization, depends on the type of business, the type of data processed, the purpose of the organization, and so on. Therefore, different types of organizations have different attractiveness and, consequently, are exposed to different levels of risk. Therefore, in the proposed model, the previously obtained probability of success (P) is, first, weighted according to the attractiveness of the organization through the following equation:

$$p = P - w \times P \tag{6}$$

where:

- ✓ P = likelihood of successfulness;
- ✓ p = weighted likelihood of successfulness;
- ✓ w = considered weight;

The value of w varies according to the attractiveness of the organization. The values considered in the proposed model are shown in *Table 1.*

Attractiveness	Very Low	Low	Medium	High	Very High
w	40%	30%	20%	10%	0%

Table 1: Weights (w) in function of the attractiveness of the organization.

Exploiting the attractiveness of the organization, it is also possible to estimate the amount of cyber-attack attempts (n) that the venture will be subjected to in a given period (i.e., taking an example of 1 year). The values considered in the proposed model, as an example of application, are shown in *Table 2*. Note that, to simplify the model, we have introduced the hypothesis that the cyber-attack attempts are not correlated on to another, so it is impossible to consider them as preparatory attacks. Simply speaking, the identical attackers will perpetrate the equal attack (i.e. without changing its characteristics according to the outcome of the previous ones), while dissimilar attackers will not communicate and / or share information.

Attractiveness	Very Low	Low	Medium	High	Very High
n	2	4	6	8	10

Table 2: Number of attack attempts per year in relation to attractiveness.

The weighted likelihood of successfulness (p) previously obtained, together with the number of attack attempts (n), is used to estimate the probability that the considered organization will suffer a successful attack in a year. In fact, it is not relevant to determine the likelihood of having at least one attack, since, once an attack is successful, or the organization notices the attack in progress, the organization will likely improve its posture and maturity, changing so the initial conditions. While there is no shortage of cases of attacks that have been successful and have not been detected for a long time, the higher the maturity, the less likely this situation is to occur.

The purpose, therefore, to estimate the probability of having exactly one successful attack, is to stop the probability calculation as soon as an attempt is successful. To do this we have chosen to exploit the characteristics of the geometric distribution.

The geometric distribution provides the statistical distribution of the exact number of failures preceding the first successful attack. Let p be the probability of success in a single attack attempt (assumed constant for the reasons explained above), the probability that the first success occurs after i-1 failures, under the aforementioned hypothesis of statistical independence between distinct attempts, is given by following expression:

$$L(i) = (1-p)^{(i-1)} \times p \tag{7}$$

In our case, the starting point is the value of n, as given in *Table 2,* as a function of the attractiveness in *Table 1.* This value can be interpreted as the maximum number of cyber-attacks suffered by the organization, for a given posture, during the year. So, summing the $L(i)$'s, $with\ i\ =\ 1, 2 \dots, n$, gives the probability L that the attack is successful after 1 or 2 or ... n attempts, at the most (such events are mutually exclusive). This probability obviously depends on the maximum number of attacks suffered by an organization, becoming higher and higher, as expected, for increasing values of n.

As already explained:

$$L = \sum_{i=1}^{n} q^{i-1} \times p \quad (8)$$

where:

- ✓ p = weighted probability of successfulness;
- ✓ $q = 1 - p$ = probability of failure;
- ✓ n = (estimated) maximum number of attacks per year.

As stated before, (7) gives as output the cumulative distribution of the probability that the considered organization will be successfully attacked during the year. In fact, assuming, for example, to have a medium appealing organization (and so, in *Table 2*, $n = 6$), this equation evaluates all the possible situations: the probability that the first attack is successful, the probability that there is a failure before the successful attack, the probability that there are two failures before the successful attack, and so on until the probability that there are $n - 1 = 5$ failures before the successful attack. *Fig. 7* shows how the likelihood of occurrence of a successful attack L changes in relation to the probability of success p and the number of attacks n. It can be seen how, fixing a value of n, the likelihood increases as p increases while, for a fixed value of p, the likelihood increases as n increases.

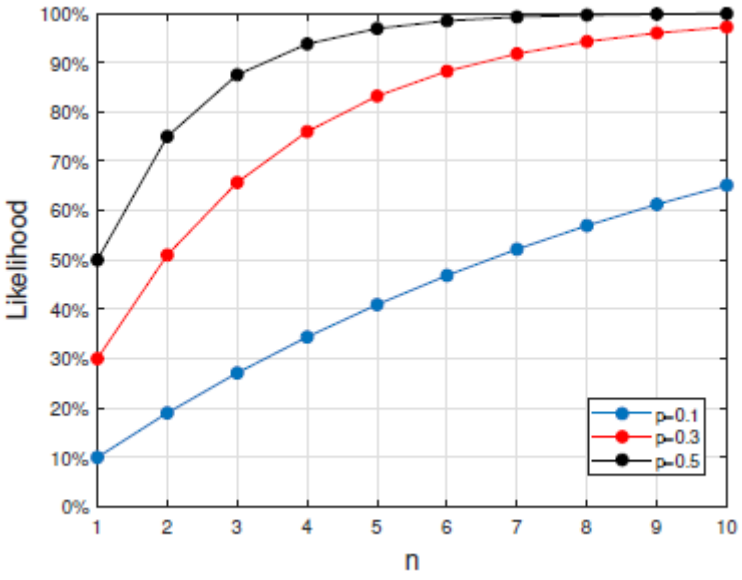

Figure 15: Variation of L in relation to p and n.

Therefore, continuing the example from *Figure 15*, we have a probability of success $(P)\ of\ 29\%$. Assuming that the considered organization has a medium attractiveness, we obtain, through (5) with $w = 20\%$ (the weight is derived from *Table 1*), that the weighted probability of success (p) is 23%. Considering the medium attractiveness of the organization and according to the values in *Table 2*, we set $n = 6$. At this point, through (7), we obtain that the probability of being successfully attacked during the year for the considered organization ($complexity\ index = 5$, $maturity\ index = 6$, $medium\ attractiveness$) is 80%. In particular, the probabilities for $i = 1, 2, \dots 6$ and the cumulative distribution are shown in *Table 3* and *Figure 13* respectively.

i	Likelihood
1	23%
2	18%
3	14%
4	11%
5	8%
6	6%

Table 3: Probability of successful attack as i varies.

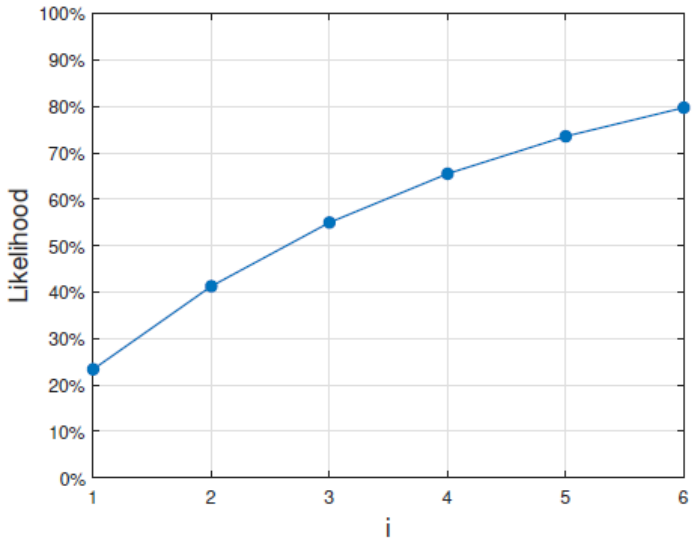

Figure 13: Cumulative distribution of one successful attack in one year as i varies.

5 Validation of the Model

After having amply explained the model proposed in the previous paragraphs, the last fundamental part of a finite model analysis turns out to be the analysis of the results: specifically, validation of the model and the evaluation of the accuracy of the outcomes. The ultimate goal would be the

characterization and validation of the model applied to the real world, and in the current thesis, our specific scope would be the implementation of the model in healthcare organizations, to obtain a greater comprehension of the phenomena and a suitable refined model *risk assessment*.

What is known by experience and by literature, is that it is relatively unpretentious to make a simulation formulating hypothesis, but it is extremely difficult to make accurate models that validate the assumptions made obtaining. Later on, several aspects related to the validation and verification of simulation models are going to be examined and analyzed. The first part of the critical analysis of the results is aimed at verifying whether the model analyzed corresponds conceptually to what was our intention or if the chosen restraints have been set correctly.

Our bibliographic research has been based on the detailed modeling of cyber-security concert and performance, with particular attention to corporate maturity and cyber-resilience, i.e. “the is the ability for organizations to prepare for, respond to and recover from cyber-attacks and security breaches” [25] as a key element of differentiation between the different organizations, basically the complexity index, in order to apply the model.

Accenture, one of the global professional services consulting companies with certified advanced capabilities in the digital, cloud and security fields, has published a 2020 Cyber-Security Report [27], alike every years, in order to help partners, clients customers and the entire cyber-ecosystem community to keep up with menaces and threats that could potentially hit their businesses, economic incomes and the entire organizations. Their analysis, based on a sample of 4644 executives ($n = 4644$), reveals that there are two distinct groups of organizations that possess differentiate cyber-security performances: leader companies, corresponding to the 17% of the sample analyzed and non-leader companies, 74% of the entire amount of executives and organizations considered. The first group, corresponding to an elite group, seemed to possess more cyber-resilience if compared to the second group obtaining significantly higher levels of performance. As the Report continues its examinations on ventures, it shows some statistical characteristic that differentiate the two different types of organizations.

CHARACTERISTICS	LEADERS (17%)	NON-LEADERS (74%)
Stop more attacks	1 in 27 attacks breach security	1 in 8 attacks breach security
Find breaches faster	88% detect breaches in less than one day	22% detect breaches in less than one day
Fix breaches faster	96% fix breaches in 15 days or less	36% fix breaches in 15 days or less
Reduce breach impact	58% of breaches have no impact	24% of breaches have no impact

Figure 14: Differentiation amongst leaders and non-leaders, Source: Accenture 2020 Report [27]

What is important to our purposes, is the differentiation between the attacks that the specific organization has correctly stopped or prevented (attacks breach security) and the total amount of attacks suffered by those specific ventures. In the first line of *Figure 14,* this discrepancy is quite visible and permits us to obtain the likelihood of successfulness of a cyber-attack, basically the result of the calculation with the logistic curve, of both leaders and non-leaders.

Leaders, the 17% of total businesses, tend to stop 1 attack over 27 suffered. Thanks to simple calculation we obtain:

$$1/27 = 0{,}037$$

Therefore, estimating the match by calculating the percentage, it is obtained:

$$0{,}037 * 100 = 3{,}7\%$$

Hence, leaders' organizations tend to have approximately the 4% of likelihood of successfulness of a cyber-attack.

$$1/8 = 0{,}125$$

Shifting to non-leaders' organization, the 74% of total businesses, they tend to stop 1 cyber-attack over an amount of 8 suffered. As before we obtain:

$$0{,}125 * 100 = 12{,}5\%$$

Henceforth, non-leaders' organizations possess the 12,5% of likelihood of successfulness of a cyber-attack, way higher than leaders, as expected.

The following step would be the exploitation of the approach amply explained in the previous chapter, calculating the likelihood of occurrence of a cyber-attack and the likelihood of successfulness of it, trying to understand how the maturity index and the attractiveness of businesses shape and affect the logistic curve. To better represent the model, the measures of asymptotes have been changed, so the lower asymptote is 0.03 for $x = 10$, the upper asymptote is 0.97 for $x = 0$. We also use a different *growth rate* such as $B = -2.$ (5)

In the case of leaders' organizations, the resulting logistic curve giving a specified set of parameters would be:

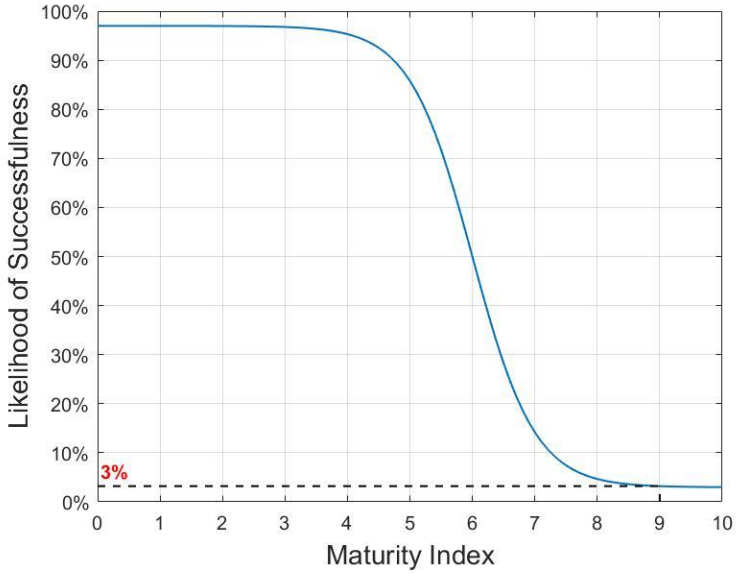

Figure 15: Likelihood of Successfulness for leaders when $m = 9$, $B = -2$, $x_0 = 6$, $att = 4$

Leaders' organizations have been modeled as follows:

- ✓ $m = 9$ as *maturity index*, thanks to the fact that leaders have advanced best security practices and subsets of controls resulting in elevated cyber-resilience;
- ✓ $att = 4$, the *attractiveness* for leaders has been set to high, since attackers possess more appealing when trying to damage those specific ventures influencing the nature of cyber-attackers and attacks their-selves;
- ✓ $x_0 = 6$ as *complexity index,* has been set to an average value, and it could not be easily modified since the complexity of the infrastructure is related to several factors, one of them related to the size of the organization, which is an unknown intrinsic parameter and therefore undying.

The logistic curve with the proposed set of parameters confirm that leaders' ventures have the 3% of *likelihood of successfulness* of a cyber-attack, similar to the percentage seen in Accenture Report [27] the sea first validation.

When changing one particular parameter, *maturity index* for a chance, trying to best fit our logistic curve, a confirmation is obtained, in fact leaders' organizations have (when *maturity index* is $m = 8$) the 4% of *likelihood of successfulness* of a cyber-attack, as it could be seen shortly after in *Figure 16:*

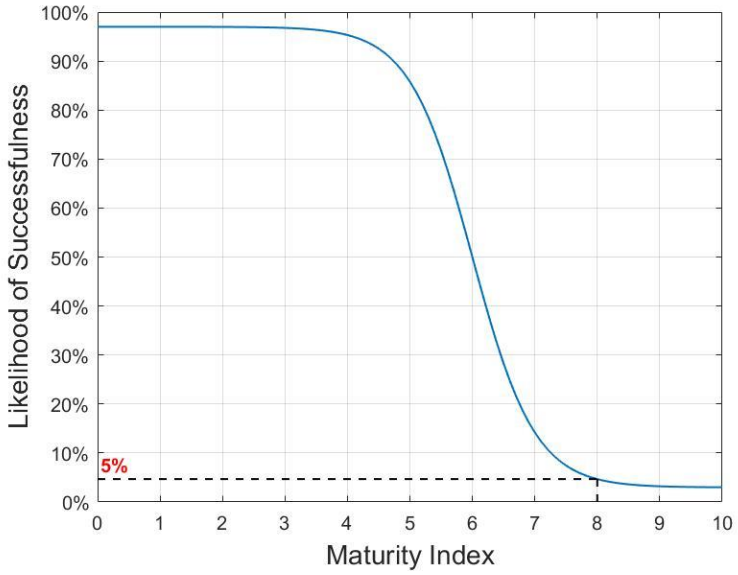

Figure 16: Likelihood of Successfulness for leaders when m=8, B=-2, $x_0 = 6$, att=4

Subsequently, we obtain the *likelihood of occurrence* of a cyber-attack for leaders' organizations, 25%, previously explained as the number of cyber-attacks suffered by the organization that would inherently and directly affect the likelihood of occurrence of suffering a successive successful attack, as:

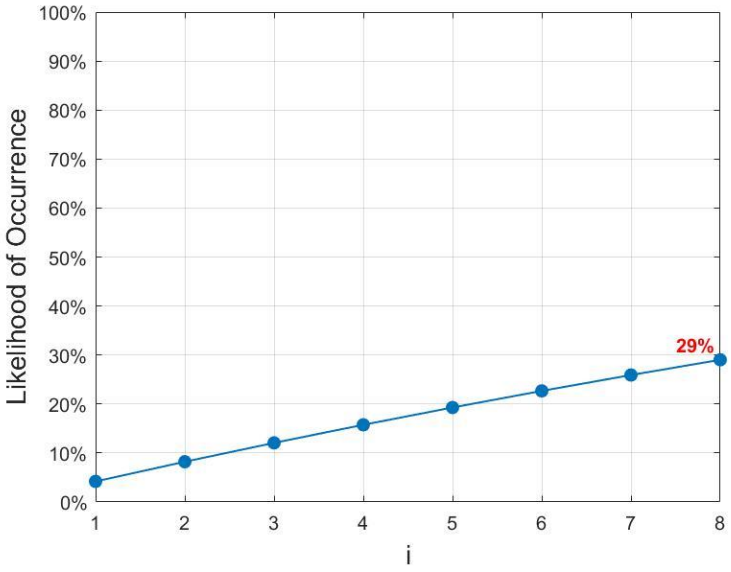

Figure 17: Likelihood of occurrence for leaders when m=8, B=-2, $x_0 = 6$, att=4, $n = 8$

How does the *likelihood of occurrence* change for $maturity\ index = 9$, considering that the other indexes and parameters remain unchanged? Let's see *Figure 18:*

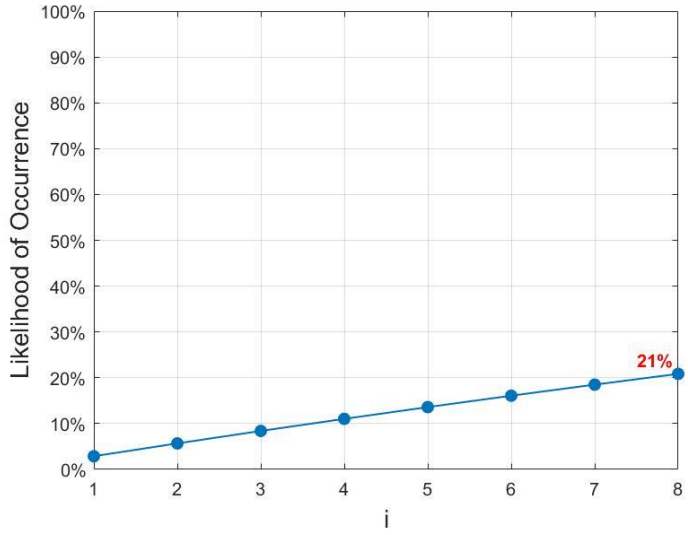

Figure 18: Likelihood of occurrence for leaders when m=9, B=-2, $x_0 = 6$, att=4, n=8

Shifting to non-leader's organizations, the shape of the logistic curves concerning *the likelihood of successfulness* and the *likelihood of occurrence* of a cyber-attack, for a given set of parameters, would be:

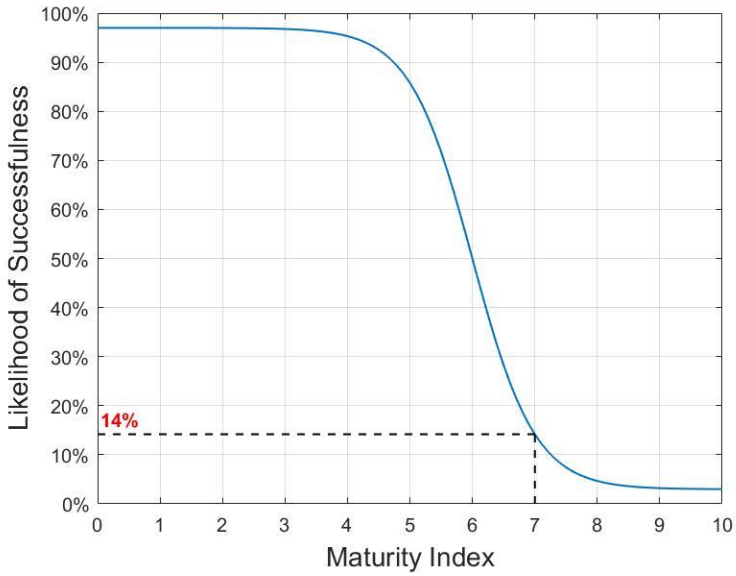

Figure 19: Likelihood of successfulness for non-leaders when m=7, B=-2, $x_0 = 6$, att=2

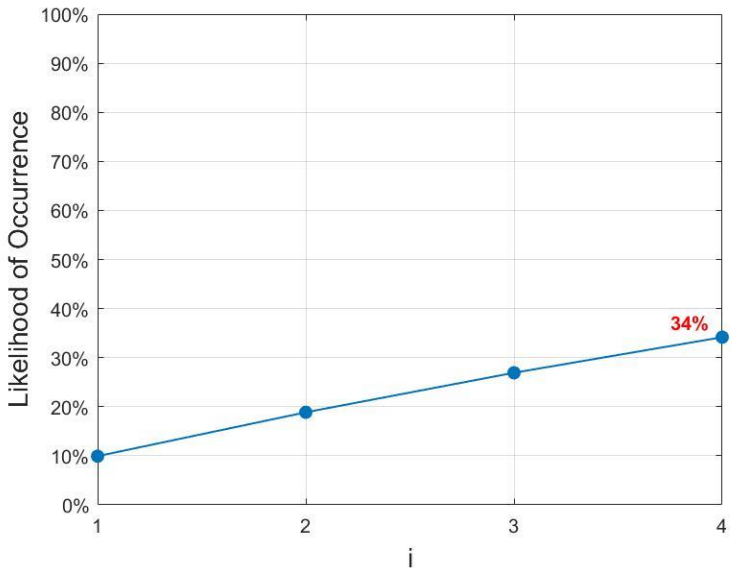

Figure 20: Likelihood of occurrence for non-leaders when $m = 7, B = -2, x_0 = 6, att = 2$

Non-Leaders' organizations have been modeled as follows:

- ✓ $m = 7$ as *maturity index*, thanks to the fact that non-leaders have less advanced best security practices and subsets of controls resulting in a diminution of cyber-resilience;
- ✓ $att = 2$, the *attractiveness* for non-leaders has been set to low, since attackers possess less appealing when trying to damage those specific ventures influencing the nature of cyber-attackers and attacks their selves;
- ✓ $x_0 = 6$ as *complexity index,* has been set to an average value, and it could not be easily modified since the complexity of the infrastructure is related to several factors, one of them related to the size of the organization, which is an unknown intrinsic parameter and therefore undying.

Our logistic curve with the second subset of parameters regarding non-leader's organization shows the 14% of *likelihood of successfulness* of a cyber-attack, very close to the percentage of *likelihood of successfulness* proposed in Accenture Report [27] improving our hypothesis.

Once the likelihood of occurrence of a cyber-attack for non-leaders' organizations have been obtained, as the 33% in this specific case, it is possible to assert that a non-leaders' venture would suffer a higher number of cyber-attack, being less cyber-resilience and achieving significantly lower levels of performance if compared to the leaders 'one, resulting in a direct affection of the *likelihood of occurrence* of suffering a successive successful attack.

Another important Report that has been analyzed, is the one proposed by Detica [28], published in partnership with the office of cyber security and information assurance in the Cabinet Office.

The purpose of the Report is to estimate the cost associated to cyber-crime, hence the economic impact, to every illegal activity associated to cyber criminals to exploit business vulnerabilities, for financial gain and to illicitly attack or gain personal and sensitive information. Their methodological assessment attempted to determine the motivations, means and opportunities presented to potential cyber-attackers. It recognizes that the nature of industrial espionage in different sectors of activity is different and has different levels of exploitation and economic impact if it is stolen. For their ideologies, cyber-criminals are more likely to target espionage organizations based on perceived size and revenue rather than the area of activity in which they operate. Based on the afore-mentioned statement, it is easy to assess that their Report exploit, and is based on the overall size of a company rather than the business sector, when talking about estimates of cost of industrial espionage, giving a differentiation as: small, medium, large business. Their differentiation is based on the datum that small businesses tend to have lesser annual costs if compared to medium and large companies.

Therefore, our aim would be the application of the new model assessment to the healthcare world. In order to have a better comprehension of the behavior of the logistic sigmoid function under different parametrization with the goal of the implementation of the proposed model assessment in healthcare companies, we archetype the model as it goes.

Since the model will be applied to the same type of organization, the medical field in our specific case, the attractiveness has been set at an appropriate value and maintained constant. From here

we started the parameterization process by changing one parameter from time to time observing the variation of the others concerning that given parameterization. Attractiveness' numerical value for healthcare organizations has been initially set to 4, $att = 4$, and initially $complexity\ index = 7$ and $maturity\ index = 7$. The first attempt has been the evaluation of the shape of the sigmoid function along with its behavior, so the likelihood of occurrence of an adverse event along with the likelihood of successfulness of the harm-attempt when the $maturity\ index$ of organizations changes.

Let's obtain the graphs with the initial constraints imposed evaluating each time their changes. The successive plots would be our initial point of evaluation.

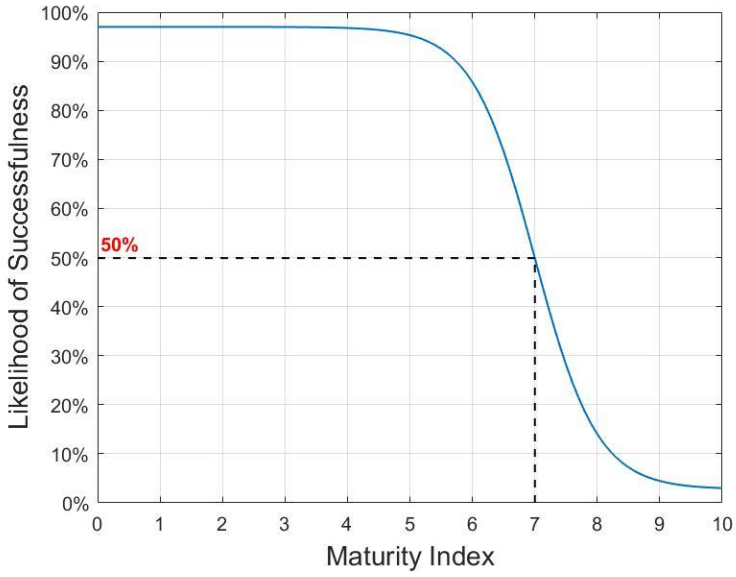

Figure 21: Likelihood of successfulness for healthcare companies when $m = 7, B = -2, x_0 = 7, att = 4$

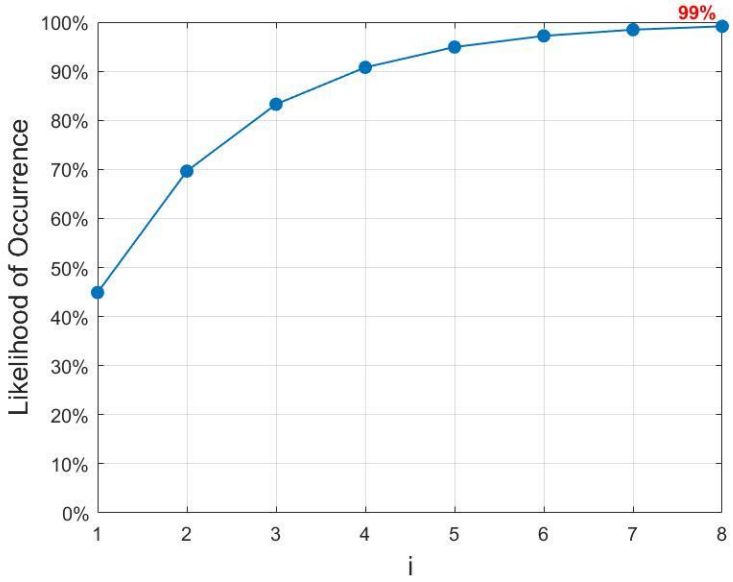

Figure 22: Likelihood of occurrence for healthcare companies when $m = 7, B = -2, x_0 = 7, att = 4$

Now, let's change for a chance $maturity\ index = 8$. The resulting plots would be:

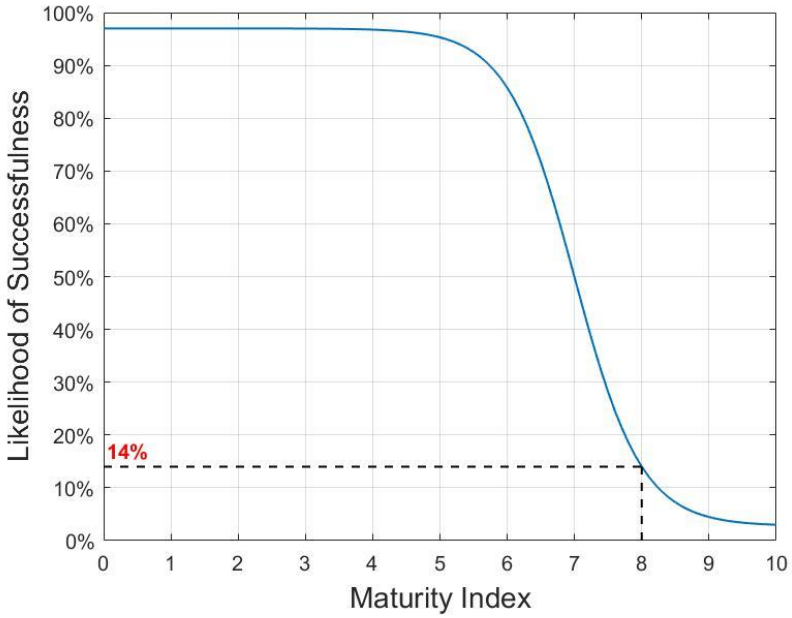

Figure 23: Likelihood of successfulness for healthcare companies when $m = 8, B = -2, x_0 = 7, att = 4$

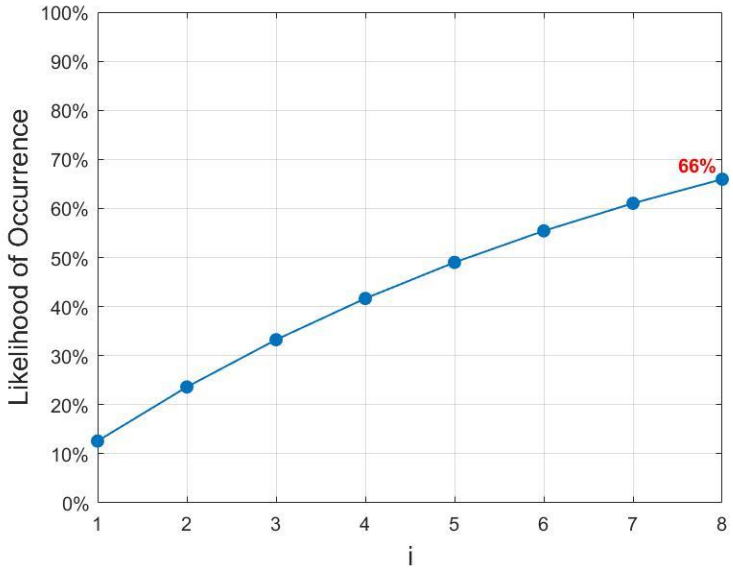

Figure 24: Likelihood of successfulness for healthcare companies when $m = 8, B = -2, x_0 = 7, att = 4$

As last attempt, we sat $maturity\ index = 9$. The resulting graphs are shown hereafter:

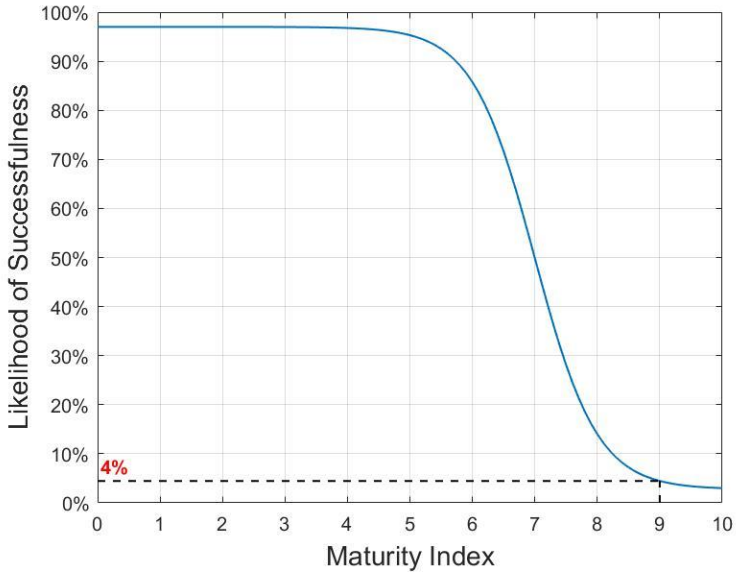

Figure 25: Likelihood of successfulness for healthcare companies when $m = 9, B = -2, x_0 = 7, att = 4$

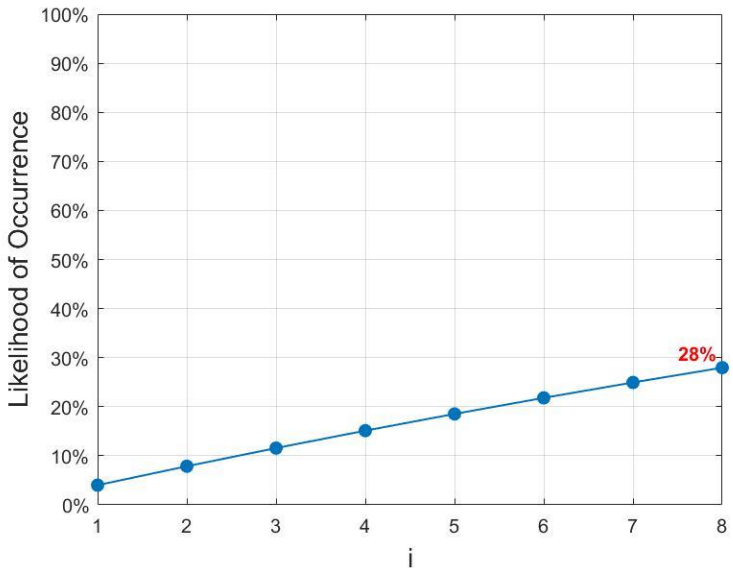

Figure 26: Likelihood of occurrence for healthcare companies when $m = 9, B = -2, x_0 = 7, att = 4$

The second attempt has been the evaluation of the shape of the sigmoid function along with its behavior, so the likelihood of occurrence of an adverse event along with the likelihood of successfulness of the harm-attempt when the $complexity\ index$ of organizations changes.

Imposing the $complexity\ index = 8$, the subsequent graphs obtained would be:

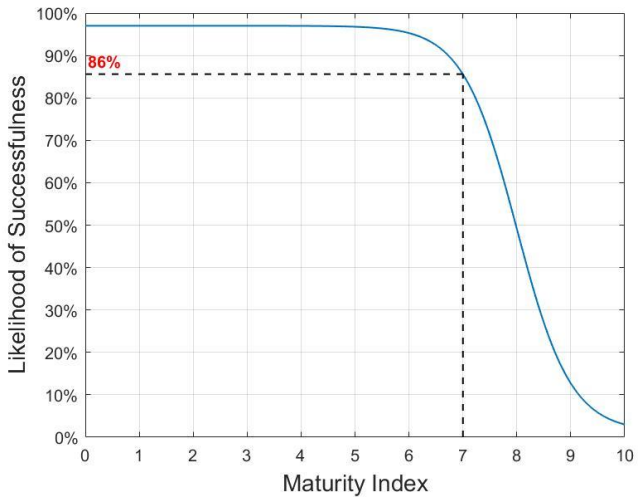

Figure 27: Likelihood of successfulness for healthcare companies when $m = 7, B = -2, x_0 = 8, att = 4$

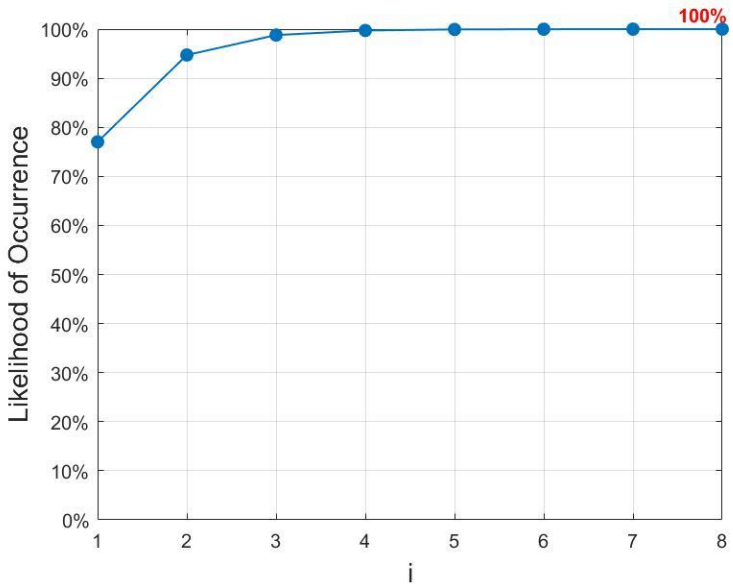

Figure 28: Likelihood of Occurrence for healthcare companies when $m = 7, B = -2, x_0 = 8, att = 4$

As last attempt, we sat $complexity\ index = 9$. The resulting graphs are shown hereafter:

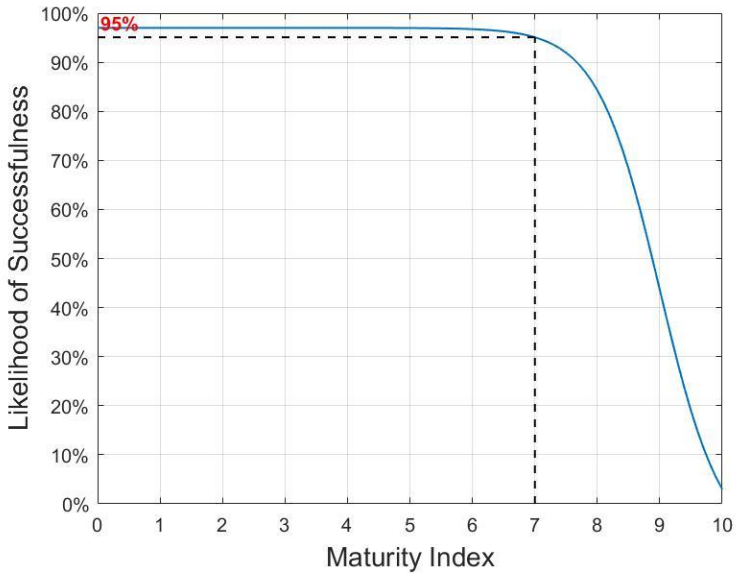

Figure 29: Likelihood of successfulness for healthcare companies when $m = 7, B = -2, x_0 = 9, att = 4$

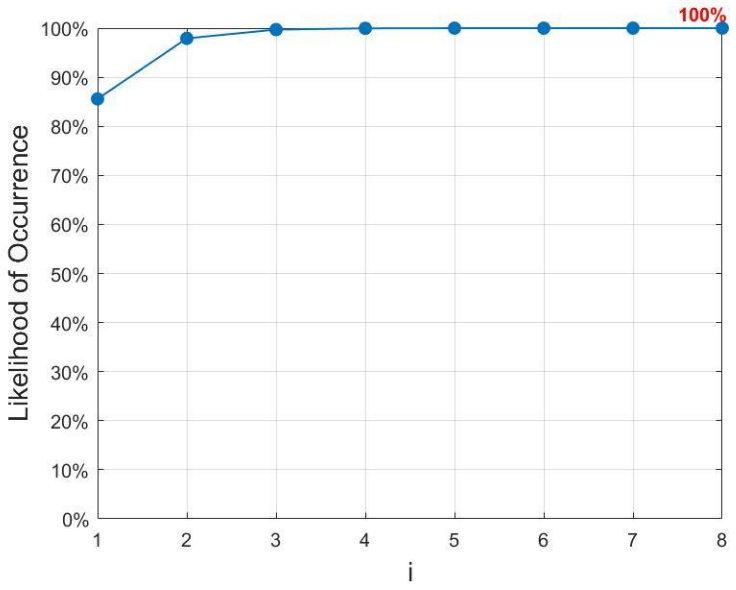

Figure 30: Likelihood of occurrence for healthcare companies when $m = 7, B = -2, x_0 = 9, att = 4$

As the proposed parametrization aims to validate the hypothesis made in Section 4, some considerations must be outlined. The sigmoid logistic curve seems to follow the hypothesis, along with the few attempts proposed in literature, which states that when $maturity\ index$ is augmented, in this case keeping the $complexity\ index$ and the $attractiveness$ unchanged, the *likelihood of occurrence of a cyber-attack* along with *the likelihood of successfulness* of it diminishes *(Figure 23-27).* Under the same premises when the $complexity\ index$ changes to a higher value, keeping the $maturity\ index$ and the $attractiveness$ unchanged, the *likelihood of occurrence of a cyber-attack* along with *the likelihood of successfulness* of it increases *(Figure 28-30).*

These results seem to agree with the insufficient notions present in the literature [26] that correlate the increase of the $complexity\ index$ to an increase in the *probability of occurrence and successful*ness of cyber-attacks. At the same time, an increase in the $maturity\ index$ results in a decrease in the probability of occurrence and successfulness of cyber-attacks, consistent with the few assumptions currently present in literature [27]. Besides, testing the new assessment model with a fabricated case study through the combination of a $complexity\ index$ and a $maturity\ index$ and through the application of a generalized logistic function, shows correspondences with the limited concepts present in the actual state of the art regarding the *risk model assessment* but, of course, failed to validate the proposed model assessment with the chosen parametrization.

6 Conclusions

In the present thesis, it has been discussed *Risk Management* as the identification, evaluation, and prioritization of risks (defined in ISO 31000 as the effect of uncertainty on objectives) followed by coordinated and economical application of resources to minimize, monitor, and control the probability or impact of unfortunate events or to maximize the realization of opportunities. Later,

we focused our attention on qualitative, semi-quantitative and, quantitative methods for risk assessment, their intrinsic difference evaluating the pros and cons of each methodology. It has been chosen to follow a quantitative method, a more highly structured method based on probability and likelihood capable of producing numerical data instead of descriptive ones thanks to structured observation using questionnaires and surveys to ensure the rigor and reproducibility of the results. We focus our attention discussing *Risk Management Frameworks* with particular regard to the National Cybersecurity Framework, a fundamental tool to support organizations that need strategies and processes aimed at protecting personal data and cyber-security posing particular attention to data protection and how GDPR has been or could be contextualized in this increasingly modern new world. It has been underlined the difficulty when addressing the topic of data protection since it requires organizations to equip themselves with suitable measures to protect the confidentiality, integrity, availability of data, as well as guaranteeing the resilience of the processes when processing the aforementioned data. Our initial aim was to present the current state of the art regarding *Risk Management Frameworks*, the aspect they considered, how they operate the difficult *Risk Management* tasks' their differences and, possible improvements. The idea underlying this first part is the implementation of *Risk Management* methodology and frameworks to address healthcare as the main scope, trying to give a standardized and generalized methodology when calculating the risk. The second part of this thesis has been focused on a healthcare case study, as Ospedali Riuniti. We amply discussed the implementation and drafting of a treatment register or data processing activities register compliant with GDPR and capable to assess the risk, testing their former tool, and proposing a new one trying to surpass the critical issue that arose. The idea was the proposal of a finite quantitative tool instead of the defective one they implemented. In particular, the lack of technological context nor information about IT infrastructure, the confused aggregation of processes and treatments prevent the correct assessment of the risk analysis due to the lack or discrepancy of data. Is in this case scenario that a new model assessment, based on a *quantitative approach*, has been proposed, trying to solve uncertainties about the current methodology giving an objective approach. The proposed model combines a maturity and a complexity index through the application of a generalized logistic

function, allows the estimate of the probability for a specific organization to be successfully attacked during a defined time. It also gives information about the actual posture of an organization, that, along with the maturity aspect would give a wide overview about the asset of the organization simplifying the process of risk assessment. It has been tried to validate the model, starting from fictious case studies found in literature, as Accenture [27] or Detica Reports [28]. Unfortunately, due to the poorness of information and case studies present in the literature, this goal has not been achieved. A parameterization has been implemented (clearly it is never unique) the parameter (or parameters) can be chosen in different ways depending on the type of curve, equation or to simplify the calculations, to see if the assumptions made in the previous chapter were valid. In general, the curve seems to be consistent with the relationships identified between the parameters (indexes) explained in 4, even if a true validation of the model has not been performed but only several tests over sets of data.

The validation of the model could be a starting point for future improvements since it is capable to relate the likelihood of success of a cyber-attack with the maturity index. Businesses should address their cybersecurity posture, refining it continuously trying to achieve a higher level of business maturity, thus decreasing the likelihood of success of an adverse event, and to be up to new possible cyber-harms and hazards this modern world is presenting us.

Bibliography

[1] E. Leverett, R. Clayton and R. Anderson, “Standardisation and certification in the 'Internet of Things',” in 16th Annual Workshop on the Economics of Information Security (WEIS), 2017

[2] P. Williams and A. J. Woodward, “Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem,” Medical Devices. Evidence and Research, vol. 8, pp. 305-316, 2015

[3] National Institute of Standards and Technology (NIST). Special Publication 800- 30 Revision 1 - Information Security

[4] Radu (Genete), Laura-Diana. (2009). Qualitative, semi-quantitative and, quantitative methods for risk assessment: Case of the financial audit. Analele Stiintifice ale Universitatii "Alexandru Ioan Cuza" din Iasi - Stiinte Economice. 56. 643-657.]

[5] Stefan Taubenberger et al. \Problem Analysis of Traditional IT - Security Risk Assessment Methods { An Experience Report from the Insurance and Auditing domain". In: Future Challenges in Security and Privacy for Academia and Industry. SEC 2011. IFIP Advances in Information and Communication Technology. Ed. by Camenisch J. et al. Vol. 354. Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-21424-0_21.

[6] GDPR and healthcare: Understanding health data and consent | Pega

[7] International Organization for Standardization (ISO). ISO/IEC 27001:2013 - Information technology | Security techniques - Information security management systems | Requirements

[8] International Organization for Standardization (ISO). ISO/IEC 27017:2015 - Information technology | Security techniques - Code of practice for information security controls based on ISO/IEC 27002 for cloud services.]

[9] International Organization for Standardization (ISO). ISO 31000:2018 – Risk management - Guidelines.

[10] https://survey.charteredaccountantsanz.com/risk_management/small-firms/context.aspx

[11] Center of Internet Security (CIS). CIS Controls v.7.1. Version 7.1.

[12] CIS-Sapienza - Research Center of Cyber Intelligence and CINI Cybersecurity National Lab - Consorzio Interuniversitario Nazionale per l'Informatica Information Security - Sapienza Università di Roma. Framework Nazionale per la Cybersecurity e la Data Protection. Version 2. Feb. 2019.

[13] Agenzia per l'Italia Digitale (AgID) - Presidenza del Consiglio dei Ministri – Area Sistemi tecnologie e sicurezza informatica - Misure minime di sicurezza ICT per le pubbliche amministrazioni. Apr. 2016.

[14] Information Systems Audit and Control Association (ISACA). COBIT 2019 (Control Objectives for Information and related Technology).

[15] Gestione del rischio/risk management: prendere decisioni aziendali sicure - IONOS

[16] F. Gradozzi, S. Leonarduzzi and G. Libertini, Privacy in regola, 2019]

[17] cybersecurityframeworks.it

[18] https://protezionedatipersonali.it/direttiva-nis-network-information-security

[19] I. Lopes and P. Oliveira, “Implementation of General Data Protection Regulation: a survey in health clinics,” IEEE, 2018

[20] Setting the Scope and Limits of a Risk Assessment (assp.org)

[21] Terje Aven, Risk assessment and risk management: Review of recent advances on their foundation, European Journal of Operational Research,Volume 253, Issue 1, 2016, Pages 1-13, ISSN 0377-2217, https://doi.org/10.1016/j.ejor.2015.12.023

[22] hendriks_MA_bms.pdf (utwente.nl)

[23] Ioannis Agrafiotis, Jason R C Nurse, Michael Goldsmith, Sadie Creese, David Upton, A taxonomy of cyber-harms: Defining the impacts of cyber-attacks and understanding how they propagate, Journal of Cybersecurity, Volume 4, Issue 1, 2018, tyy006, https://doi.org/10.1093/cybsec/tyy006

[24] Cambridge Online Dictionaries

[25] Cyber resilience - GOV.UK (www.gov.uk)],

[26] Increasing Complexity Creates Challenges for Risk Management | ERM - Enterprise Risk Management Initiative | North Carolina State Poole College of Management (ncsu.edu)

[27] https://www.accenture.com/_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf

[28] Detica - the-cost-of-cyber-crime-full-report.pdf (publishing.service.gov.uk)