POLITECNICO DI TORINO

Master's Degree in Computer Engineering

Master's Degree Thesis

Towards Automated Information Gathering and Processing for Cyber Risk Assessment

Supervisor:

Prof. Cataldo Basile

Candidate:

Gabriele Gatti

Academic Year 2021/2022

Torino

Abstract

During the past decades, technology quickly took a predominant role within the architecture of organizations and enterprises of all sizes and purposes. From public relations to business functions, information systems are employed to simplify tasks and optimize workloads. However, with great benefits also come significant disadvantages: exposure to cyber threats and cyber risk. In a world where cyber attacks could easily disrupt companies and damage stakeholders, the precise evaluation of the cyber risk factor is crucial, both for companies desiring to mitigate their risk and for insurance providers. The risk analysis process is costly and time-consuming, requiring particular efforts for information gathering and the consequent processing of it. Starting from such premises, the work in this Thesis aims to explore diverse approaches to collecting risk-related information from multiple sources. Particular attention is given to the offensive point of view, the attacker's perspective with its techniques, tactics, and procedures for gathering information about their targets.

The research started with analyzing industry-level standards for cyber security and cyber risk management to extrapolate what information can be considered significant under the cyber risk assessment perspective and how such knowledge can be collected. The focus is then shifted to the attacker's point of view, and different offensive tools are evaluated to understand the risk-related information they can produce. Additionally, different strategies for enhancing such data are studied, producing interesting results in the scope of data classification. As a final result, an automated framework for cyber risk knowledge collection, processing, and enrichment is built upon the outcomes of the research's previous portions. By employing forward reasoning on starter chunks of information, the framework can programmatically collect new data based on previously defined rules and produce novel knowledge based on the gathered one. Lastly, ideas for further work are presented through extensions of the framework with new collection techniques, additional sources of information, and additional machine learning-assisted methods for data enrichment.

Acknowledgments

This Thesis represents both the end and the beginning of chapters of my life. It signifies the end of a path that lasted five years and helped me develop as an engineer and, more importantly, as a human being. All the people around me during all these years participated and helped me in my growth, making me who I am now. To my mom, and my dad, who assisted me and supported me in all my choices without ever doubting; to my brother, whose opinions always helped and reassured me when feeling unsure; to my sister, her husband, and my little nieces, that made me feel at home every time I needed; to all my friends, for never failing to make me laugh; and to Claudia, whose love and affection brightened up even my cloudiest days. To my supervisor, Prof. Basile, because without his support and experience this document wouldn't exist in the first place. To those people, I say "Grazie!" from the bottom of my heart, hoping that the chapter now beginning will hold many more opportunities to thank you again.

Ga

Table of Contents

List of Figures

List of Tables

Acronyms

API

Application Programming Interface

ATT&CK

Adversarial Tactics, Techniques & Common Knowledge

CIA

Confidentiality, Integrity and Availability

CLI

Command Line Interface

CLIPS

C Language Integrated Production System

CSRF

Cross-site Request Forgery

CVE

Common Vulnerabilities and Exposures

CVSS

Common Vulnerability Scoring System

CWE

Common Weakness Enumeration

DORA

Digital Operational Resilience Act

ECB

European Central Bank

EPSS

Exploit Prediction Scoring System

FN

False Negatives

FP

False Positives

HTTP

Hypertext Transfer Protocol

ICT

Information and Communication Technologies

IEC

International Electrotechnical Commission

ISO

International Organization for Standardization

IP

Internet Protocol

IT

Information Technology

LHS

Left Hand Side

NIST

National Institute of Standards and Technology

NLP

Natural Language Processing

NMAP

Network Mapper

OSINT

Open Source Intelligence

RHS

Right Hand Side

SLA

Service Level Agreement

SP

Special Publications

TIBER-EU

Threat Intelligence-based Ethical Red Teaming

TCP

Transmission Control Protocol

TLPT

Threat Led Penetration Testing

TN

True Negatives

TP

True Positives

TTPs

Techniques, Tactics and Procedures

UDP

User Datagram Protocol

URL

Uniform Resource Locator

XSS

Cross-site Scripting

Chapter 1

Introduction

The Thesis describes the work related to exploring techniques and methodologies that allow for automatic extraction of information from various sources, publicly available and, when possible, company provided. The information gathered will then be aimed at estimating a company's risk profile, seen from an IT security and cyber security point of view, a measure generally known as Cyber Risk.

Correctly evaluating the cyber risk profile of a company is an extremely important task: having a defined quantity that represents a company's exposure to the threats that may bring disruption to the business, not only guides the enterprise in the development of strategies to defend itself but also serves a fundamental role for third parties providing services such as insurance products, allowing those entities to customize rates with respect to the customer and its condition.

Despite the importance and usefulness of the measure, its estimation is not trivial and presents different challenges that this thesis aims to overcome and simplify, posing the basis for a streamlined process of cyber risk estimation.

The primary issue that will be tackled in the following chapters is the gathering, enrichment, and elaboration of information needed for the derivation of such measure, with a particular focus on how the process can be automated, reducing as much as possible the need for human interaction in every one of its steps. The most important reasons behind the need for automation can be found in two cardinal points:

- Efficiency of the process (compared to the manual collection);
- Reliability of the gathered data.

While the first point can be explained in a pure sense of manpower needed to fulfill the required task; the second finds its justification in the tendency of subjects to underestimate the importance of an accurate cyber risk evaluation process, which leads to incomplete or even erroneous data.

The initial phases of the work consisted of an in-depth analysis of the current literature on cyber risk, mostly comprised of enterprise standards defining the key concepts, processes, and best practices developed for the optimal assessment and consequent management of a cyber risk profile.

The examination of such documents allowed for the extrapolation of significant information considered to be useful during a risk estimation procedure. Eventually, new sources of information are also identified from the standards: vulnerability assessment activities were recognized as possible starting points for new data.

Once information is identified and, when needed, classified, the research focuses on establishing how it can be gathered and enriched in order to build a knowledge base aimed at modeling the context in which the target of the risk evaluation is situated. Simpler data collection techniques mostly rely on open source tools such as NMAP[1] or theHarvester[2], while more complex and targeted system analysis leverage specialized vulnerability assessment tools; other more traditional approaches towards information gathering such as the administration of questionnaires to enterprises has been considered but not validated. Concerning instead the enrichment of gathered data, an exploration of novel practices has been undergone, in particular, the branch of Machine Learning named NLP produced interesting results in the scope of categorization.

[1] https://nmap.org/

[2] https://github.com/laramies/theHarvester

As a final result of the research process, all the knowledge about information, its sources, and the techniques for processing it that have been learned from the previous phases are therefore channeled into a more complex system, namely an expert system[3], that is able to autonomously gather and process information relevant for risk assessment purposes from a given target. The approach for developing this system is that of forward reasoning applied to bits of information represented as facts that have been modeled from the knowledge acquired during the study; following a set of predefined rules, derived as well from such knowledge, the system is able to increase the volume of the risk-relevant data by deriving new information from the previously gathered, in some cases resorting to internal enrichment processes, while in other exploiting external sources of information.

Chapter 2 provides the reader the background about risk management and risk assessment, with a particular focus on the concepts and best practices presented by industry-level standards adopted by enterprises and organizations worldwide.

Chapter 3 presents a definition of the problem studied and the constraints and limitations introduced by the environment in which the study was conducted.

Chapter 4 begins the process of exploration of possible sources of information by selecting the key concepts of the standards that comprise the theoretical background, thus producing refined, high-level concepts that could be well suited for business surveys.

Chapter 5 shifts the focus to new possible sources of information that derive from a different approach: the vulnerability assessment. Software tools developed for this purpose are evaluated and provide an insight into what information is obtainable from them. Additionally, techniques for information enrichment through the employment of machine learning have been explored.

Chapter 6 contains the description of the final product obtained from the knowledge acquired in the previous chapters, in the form of the framework for risk-relevant information gathering that is based on an autonomous expert

[3] `https://en.wikipedia.org/wiki/Expert_system`

system. The system's structure is presented, as well as the data models and the rules employed.

Finally, Chapter 7 presents the conclusions derived from the work of this thesis, providing different ideas for future work in the form of expansion of the framework.

Chapter 2

Background

After an accurate definition of risk management's focuses, this chapter will present the enterprise standards taken into consideration, highlighting similarities and common cardinal themes when possible. From the documents studied, the importance of Vulnerability Assessment for cyber risk estimation will be emphasized, leading to a more comprehensive definition of such a process as well as an excursus on a set of tools to achieve this purpose and their outputs. Lastly, contextual information about the other instruments employed during processing the gathered data is provided, with a particular focus on the language used for the framework that has been developed to enclose and automate the comprehensive process.

2.1 Cyber Risk Management

Risk management in IT systems is a complex, macroscopic process defining those business practices and activities aimed at keeping the company secure and avoiding or at least minimizing the consequences of malicious actions against the enterprise's assets. According to the NIST Special Publication 800-39 [1], the de facto standard for information security risk management definition and guidelines, risk management is comprised of four main phases: (**i**) frame risk (i.e., establish the context for risk-based decisions); (**ii**) assess risk; (**iii**) respond to risk once determined; and (**iv**) monitor risk on an ongoing basis. Summarizing from the document:

(i) **Framing**
The first component of risk management aims at defining how enterprises establish a risk context, in other words, how they describe the environment that surrounds risk-based decisions. The purpose of the risk framing component is to produce a risk management strategy that addresses how organizations intend to assess, respond, and monitor risk, allowing a transparent manifestation of the risk perceptions that organizations employ on a day-to-day basis for making investment and operational decisions. The risk frame functions as a foundation for risk management and poses boundaries and constraints inside the organization for risk-based decisions. For the establishment of a realistic and credible risk frame, enterprises are required to identify: risk assumptions (e.g., assumptions about the threats, vulnerabilities, impacts, and likelihood of occurrence) that affect the assessment, response and monitoring of risk; risk constraints, meaning the limitations affecting the components of risk management; risk tolerances that represent the levels, types, and degree of risk uncertainty that are acceptable; priorities and trade-offs (e.g., the relative importance of business functions, compromises between different types of risk affecting the organizations and time frames in which risk needs to be addressed).

(ii) **Assessing**
The second component of risk management defines how organizations assess risk in relation to the context established with the risk frame. Risk assessment aims to identify: direct and indirect threats to organizations; internal and external vulnerabilities affecting the organizations; the damage or impact that may affect organizations as a result of threats exploiting vulnerabilities; and the likelihood that such harm will actually occur. The end result is an evaluation of risk, that is, the degree of harm and likelihood of occurrence of said harm. It is also important for organizations to evaluate and identify a series of parameters supporting risk assessment: the tools, techniques, and methodologies used to assess risk; the assumptions related to risk assessments; the constraints that may affect risk assessments; roles and responsibilities inside the risk assessment process; the frequency of risk assessments.

(iii) Responding

After the risk has been assessed, organizations need to address how to respond to risk based on the results obtained from the assessment; the purpose of the risk response component is to provide a consistent and organization-wide response to risk, coherently with the context evaluated by the risk frame. A response strategy is established by: developing, evaluating, and determining alternative paths of action for responding to risk, consistently with the risk tolerances that have been identified by the risk frame; and implementing adequate risk responses based on the chosen alternative course of action. The possible risk responses are accepting, mitigating, avoiding, sharing, or transferring risk. Lastly, organizations also identify the tools, techniques, and methodologies used to develop, evaluate, and communicate alternative courses of action for risk responding.

(iv) Monitoring

The last component of risk management describes guidelines for monitoring over time the behaviour of risk, in relation to the results of the other components. Risk monitoring serves different purposes: verifying the correct implementation of risk response measures, assuring that the relative prerequisites are satisfied; evaluating the effectiveness of said measures; and identifying changes to the information systems' infrastructure, or more in general, to the business environment that could impact risk.

While mainly Assessing **(ii)** directly shapes the work of this thesis by introducing a well-defined methodology for information gathering, as well as recommendations on how to measure the severity of such data in the context of risk analysis; also the remaining components determine interesting insights on new sources of intelligence useful for evaluating enterprise practices, enriching the knowledge base on which risk can be constructed from.

2.2 Security Standards

As clearly highlighted in the previous paragraphs, an extensive study of the standards correlated to the topics of risk management and assessment, as

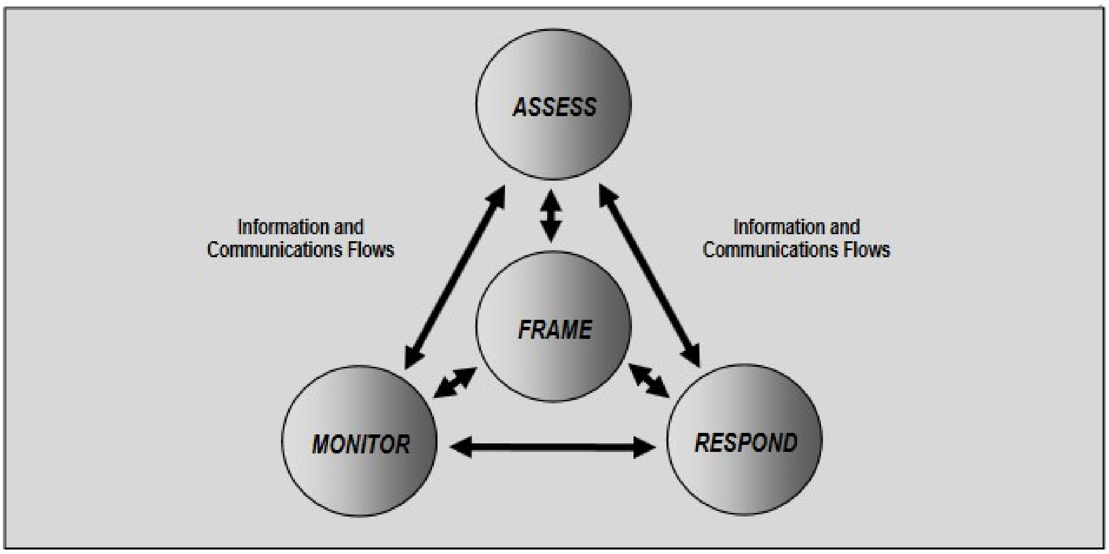

ASSESS
Information and Communications Flows
Information and Communications Flows
FRAME
MONITOR
RESPOND

Figure 2.1: Interaction between components of the risk management process [1].

well as defined best practices, is required before proceeding further in the analysis. In the following sections, the documents taken into consideration are presented.

2.2.1 NIST SP 800 Series

Created in 1990, the documents belonging to the National Institute of Standards and Technology (NIST) Special Publications series 800 present information of interest among the cyber security community. The series comprises guidelines, recommendations, technical specifications, and annual reports of NIST's cybersecurity activities. The next subsections contain an overview of cardinal concepts of the publications mostly correlated with cyber risk management, assessment, and analysis.

NIST SP 800-39

This publication [1], as already introduced in Section 2.1, thoroughly defines the concept of Cyber Risk Management and its activities.

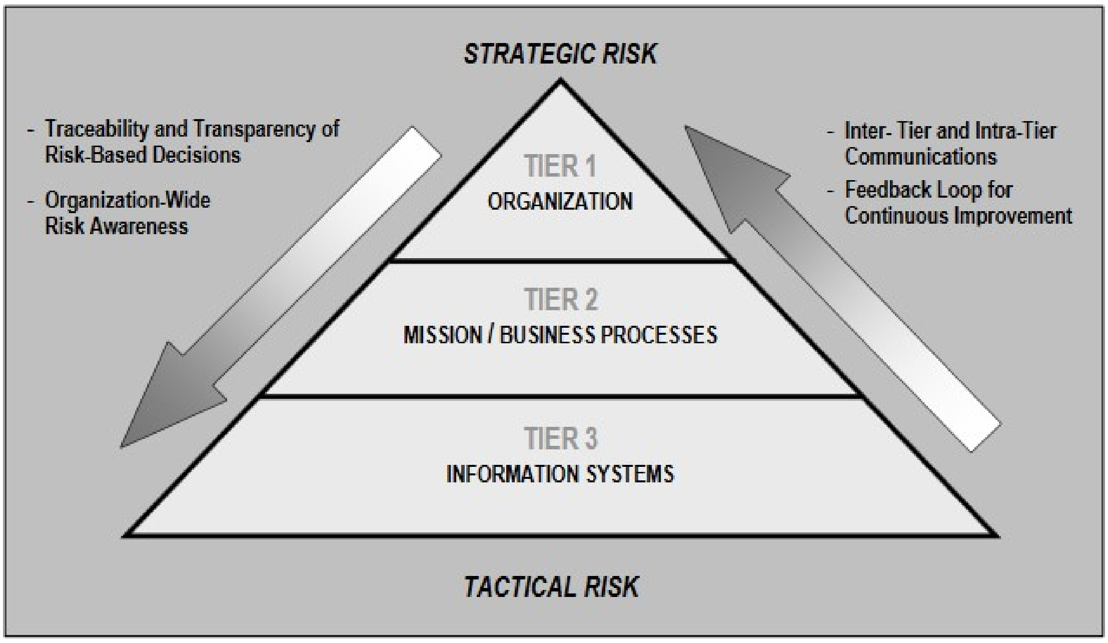

STRATEGIC RISK
- Traceability and Transparency of Risk-Based Decisions
- Organization-Wide Risk Awareness
TIER 1
ORGANIZATION
TIER 2
MISSION / BUSINESS PROCESSES
TIER 3
INFORMATION SYSTEMS
- Inter- Tier and Intra-Tier Communications
- Feedback Loop for Continuous Improvement
TACTICAL RISK

Figure 2.2: Enterprise levels of perspective for risk management [1].

Another cardinal aspect of cyber risk management yet to be discussed and strongly emphasized by the standard is the multi-tiered nature of the process. It is, in fact, important to denote how risk has to be addressed under different perspectives depending on the context inside the enterprise, specifically: organization level, mission/business process level, and information system level.

At tier 1, risk is addressed from an organizational point of view; the responsibilities in this context consist of the establishment and implementation of governance structures that introduce in the organization an oversight of the activities related to risk management, this includes the creation of the Risk Executive, the reference figure for risk inside the enterprise; the definition of a risk management strategy which regulates major decisions concerning risk; and lastly the definition of an investment strategy aiming at improving resources and security for information handling. As a consequence of this approach, tier 1 also implements risk framing, the first component of the risk management process. The decisions taken at the highest level determine the context for all the risk-related activities carried out at the lower ones.

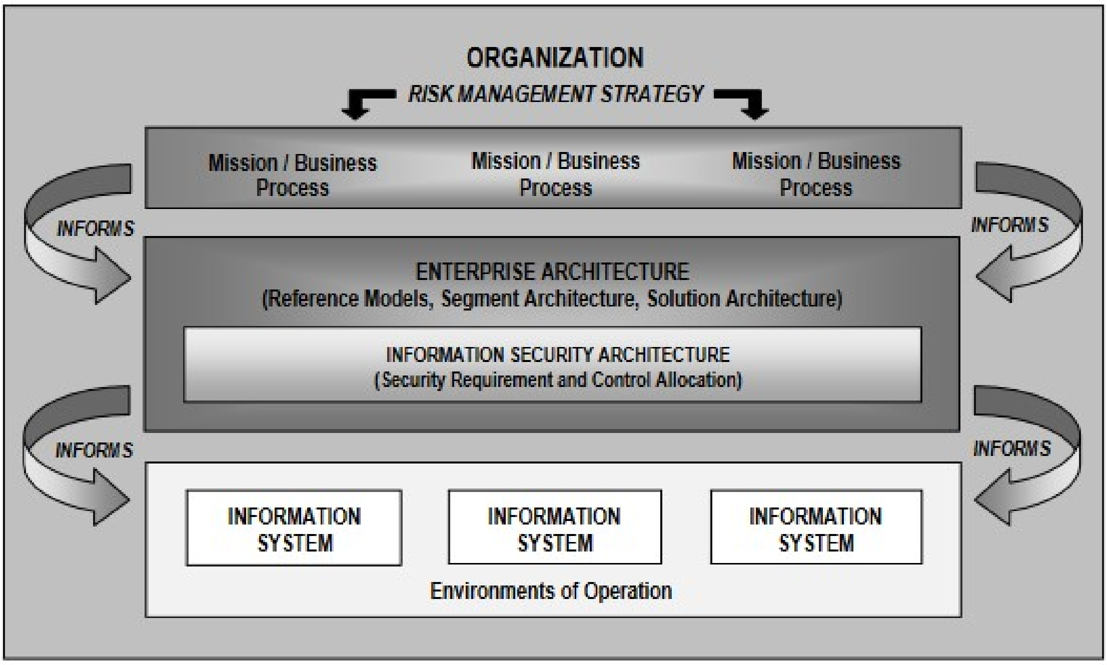

ORGANIZATION
RISK MANAGEMENT STRATEGY
Mission / Business Process
Mission / Business Process
Mission / Business Process
INFORMS
INFORMS
ENTERPRISE ARCHITECTURE
(Reference Models, Segment Architecture, Solution Architecture)
INFORMATION SECURITY ARCHITECTURE
(Security Requirement and Control Allocation)
INFORMS
INFORMS
INFORMATION SYSTEM
INFORMATION SYSTEM
INFORMATION SYSTEM
Environments of Operation

Figure 2.3: Integration of security requirements into the enterprise architecture [1].

Part of the said context is the risk management strategy, one of the key outputs of risk framing; based on priorities, assumptions, constraints, and risk tolerances identified during the process, it defines the enterprise's behaviour towards investment and operational decisions. Another integral part of framing activities is the definition of the aforementioned risk tolerance, it is the degree of uncertainty or level of risk that is acceptable to the organization; such measure clearly plays a key role in how the lower tier will approach the design phase of the business processes.

Through the developed business investment strategy, tier 1 again directly influences the prioritization of the processes developed at the lower level, this not only includes and defines the business architecture but also firmly affects those other activities strictly linked with cyber risk, identifying, for example, which assets require a higher level of protection, which as shown in the following chapters, constitutes a crucial element for risk assessment.

Tier 2 handles risk under the perspective of the business (or mission) process by introducing the concept of a risk-aware business process that can be defined as one that explicitly takes into account the potential risk that it could cause to the organization if implemented in the architecture. Thus, the duties of this level in the risk management context can be summarized in the design, development, and implementation of this class of processes.

Tier 2's work results in the creation of a well-defined enterprise architecture focused on the strategic goals and objectives of the organization that at the same time embeds all the security and risk-related decisions that have been received via the risk framing process performed at the upper level. As shown in Figure 2.3 the integration of such information in the business processes thus determines the development of a portion of the enterprise architecture entirely dedicated to the security aspects of the company, namely the security requirements and the security controls allocation inside the architecture; this part, that the document calls by the name of Information Security Architecture, undoubtedly plays a fundamental role in providing risk reduction to the organization, when the specified requirements are correctly implemented.

Tier 3 is the lowest level of perspective in the framework; from this point of view, risk related to the information systems is addressed based on the context, decisions, and activities provided by the upper tiers. Among the responsibilities of this level, the predominant ones are the categorization of information systems, the allocation of security controls to said systems and their working environments consistently with the Information System Architecture, and lastly, managing the selection, implementation, assessment, authorization, and ongoing monitoring of allocated security controls.

People in charge of this tier must take day-to-day operational risk-based decisions with a finer graded scope since they may be targeting individual information systems following the guidelines received above. This behaviour allows for a continuous feedback loop to the upper tiers that could possibly lead to enterprise architecture changes or risk tolerances adjustments; the causes triggering those events can be identified, for example, in discovering new vulnerabilities affecting organizational information systems.

The last two concepts described in the publication are closely related among them: trust and trustworthiness.

Trust can be defined as the belief that an entity will behave in a predictable manner in specified circumstances. The entity may not be limited to human beings but could also be identified as a process, as an object, or, more in general, as any combination of such elements. Trust, while being fundamentally a subjective determination, can still be based on objective and unambiguous evidence such as historical records or the results of information technology product testing and evaluation. Subjective belief, level of comfort, and experience may supplement (or even replace) objective evidence, or substitute for such evidence when it is unavailable.

On the other hand, trustworthiness is an attribute of a person or organization that provides confidence to others of the qualifications, capabilities, and reliability of that entity to perform specific tasks and fulfill assigned responsibilities. It is important to notice that this characteristic could also be applied to information technology products and systems; therefore, it is clear that the level of trustworthiness of the information systems deployed must be adequate to the maximum acceptable level of risk considered by the organization for that specific portion of the business. Failure to respect said constraint can potentially lead to a risk higher than the level accounted for.

NIST SP 800-30

This publication [2], titled "Guide for Conducting Risk Assessments" expands the work started in the previous document by providing a comprehensive review and definition of the Risk Assessment process, the most important part of the risk management framework. One of the purposes behind risk assessment can be found in addressing the potential negative consequences deriving from the operation of information systems and the information that those systems are processing, exchanging, and storing. But the risk is not only limited to information systems or even to the more general infrastructure; thus the risk assessment process aims at identifying all the possible surfaces of attack, including, but not limited to human resources, business processes, and so on. Once the risk assessment is performed, it can also support a wide variety of risk-based decisions and activities, such as the development of the information security architecture or the design and implementation of security solutions for the information systems. One key aspect of risk assessment is the fact that it is based on preconditions that tend to change over time. Hence its validity has to be limited to a certain period of time, requiring periodic repetition of the process.

Before introducing the process of risk assessment, it is crucial to define the key concepts of risk, starting from risk itself. Risk is a measure, it serves the purpose of measuring the extent to which an entity is threatened by a potential circumstance or event. Additionally, it is also possible to indicate risk as a function of two factors related to this event: the likelihood of the event occurring and the negative impact produced by the occurrence of that event. Risk can appear at every level of the organizational pyramid, for example, damage to the image or reputation of the organization or financial loss are forms of risk at Tier 1; inability to successfully execute a specific business process at Tier 2; or the resources expended in responding to an information system incident at Tier 3. With such a premise, cyber risk assessment can be explained as the process responsible for identifying, estimating, and prioritizing risks associated with the cyber security components of the enterprise.

Correctly assessing risk also requires the characterization of a well-structured methodology; the components forming such methodology are tailored over the assumptions derived from the enterprise environment of the risk assessment process. Hence those components and the whole methodology can be considered an indirect product of the risk framing process.

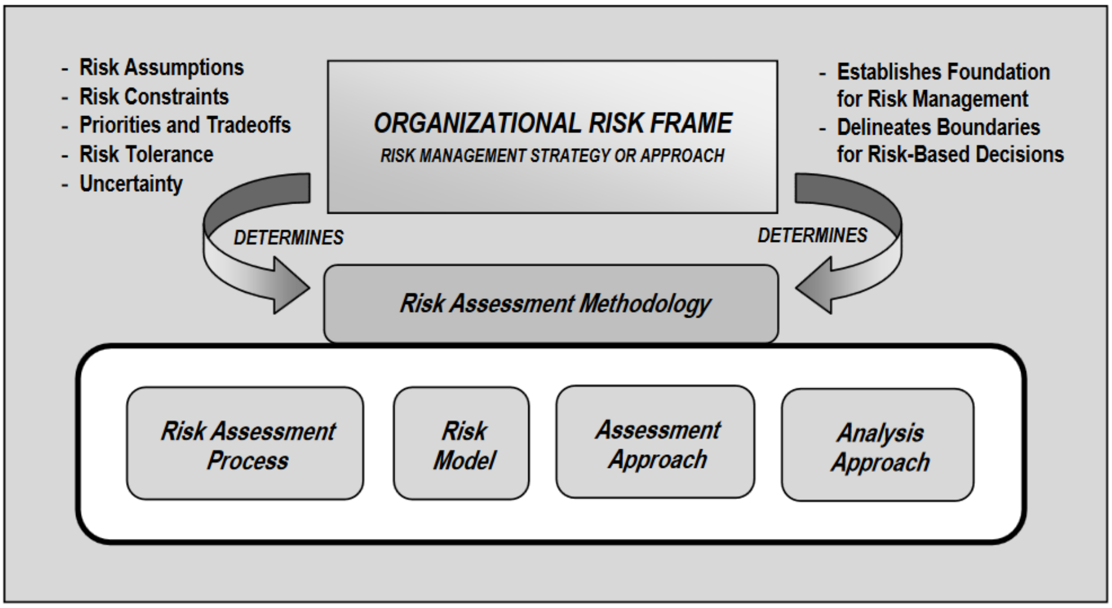

- Risk Assumptions
- Risk Constraints
- Priorities and Tradeoffs
- Risk Tolerance
- Uncertainty
ORGANIZATIONAL RISK FRAME
RISK MANAGEMENT STRATEGY OR APPROACH
- Establishes Foundation for Risk Management
- Delineates Boundaries for Risk-Based Decisions
DETERMINES
DETERMINES
Risk Assessment Methodology
Risk Assessment Process
Risk Model
Assessment Approach
Analysis Approach

Figure 2.4: Risk Assessment Methodology: relationship with risk framing and components [2].

Figure 2.4 also shows the components of the methodology; specifically, they are the assessment process which will be described in the next paragraphs; a risk model responsible for defining key terms, risk factors that can be assessed and their relationships; the assessment approach which specifies the range of values that the modeled risk factors can assume (being for example quantitative, qualitative or semi-qualitative measures) and also how those values can be functionally combined to produce an evaluation of risk; and lastly the analysis approach which describes how the modeled factors are identified and analyzed, in other words, the analysis approach determines the point of view from which the risk assessment is conducted (e.g., threat-oriented, asset/impact-oriented, or vulnerability-oriented). It is implicit that organizations can decide whether to use one or multiple methodologies for assessing risk depending on the enterprise environment.

Going further into the concept of the risk model, it is possible to notice the introduction of risk factors as the model's inputs. Typical risk factors may be listed as follows:

1. **Threats**

The term threat refers to any entity, circumstance, or event that has the potential to negatively impact the enterprise assets, individuals, and related organizations; in the context of information security, the damaging event usually happens through the organization's information systems and can be identified as unauthorized access, destruction, disclosure, modification of information, and/or denial of service. Threats can be further decomposed into threat events and threat sources, where the second causes the first.

With regard to threats, risk models differ in the level of detail and complexity to which the aforementioned characteristics are described.

2. **Vulnerabilities**
A vulnerability is a weakness that affects a portion of the organization's architecture, that being an information system, security procedures, internal controls, or any other implementation that can be exploited by threat sources. In the context of information systems, vulnerabilities can be derived from missing or weak security controls and emerge over time with the evolution of business functions and environmental changes such as the discovery of new technologies that could render certain security controls inadequate. This aspect of vulnerabilities underlines the importance of maintaining up-to-date risk assessment. Nevertheless, vulnerabilities are not only limited to information systems: organizational structures may lack adequate risk management strategies or produce inconsistent decisions; relationships with third parties may bring vulnerabilities in the form of centralized dependencies on technologies and/or resources; poorly designed business processes may lack risk aware components leading to higher exposures to risk, and poor enterprise architecture designs may undermine resiliency of the deployed information systems despite the security assumptions of those systems.

   In addition to vulnerabilities, organizations must also consider the so-called predisposing conditions, which can be defined as those characteristics able to affect (i.e., increase or decrease) the likelihood of threat events resulting in adverse impact on the organization. Those conditions may introduce new vulnerabilities that need to be accounted for.

3. **Likelihood**
The likelihood, intended as "likelihood of occurrence" is a risk factor derived from the probability analysis that a threat can take advantage of

one or more vulnerabilities. Nevertheless, it should be pointed out that this factor also conglomerates a likelihood of impact, which represents a measure of the probability that after the initiation of such an adversarial event, it also produces negative effects on the organization.

There are two approaches to the assessment of the likelihood of occurrence: when assessing adversarial threats (i.e. entities willingly aiming at attacking and potentially damaging the organization), the measure is usually based on the adversary's intents, capabilities, and targeting (meaning the level of dedication and persistence into attacking a given organization); in the case of non-adversarial threats instead, the assessment typically relies on historical evidence or empirical data. As with many other risk factors, the likelihood also needs to be bound to a certain time frame with all the derived consequences.

4. **Impact**
Impact describes the level of magnitude of the damage that can result from adversarial actions against the organization. Those actions can be identified in the following: unauthorized disclosure of information, unauthorized modification of information, unauthorized destruction of information, loss of information, or information system availability. In order to accurately define a reproducible process for assessment of impact, organizations are required to explicitly state the process used for such determinations, the assumptions behind it and the sources of information, and the methods used for extraction. To further increase the clarity of the process, an organization can also make explicit the priority and values followed for identifying high-value assets and the potential adverse impacts to organizational stakeholders.

   Particular importance lies in the fact that the risk tolerance assumptions produced in the risk framing phase may affect threat events with below-threshold impact values, stating that no further analysis is required.

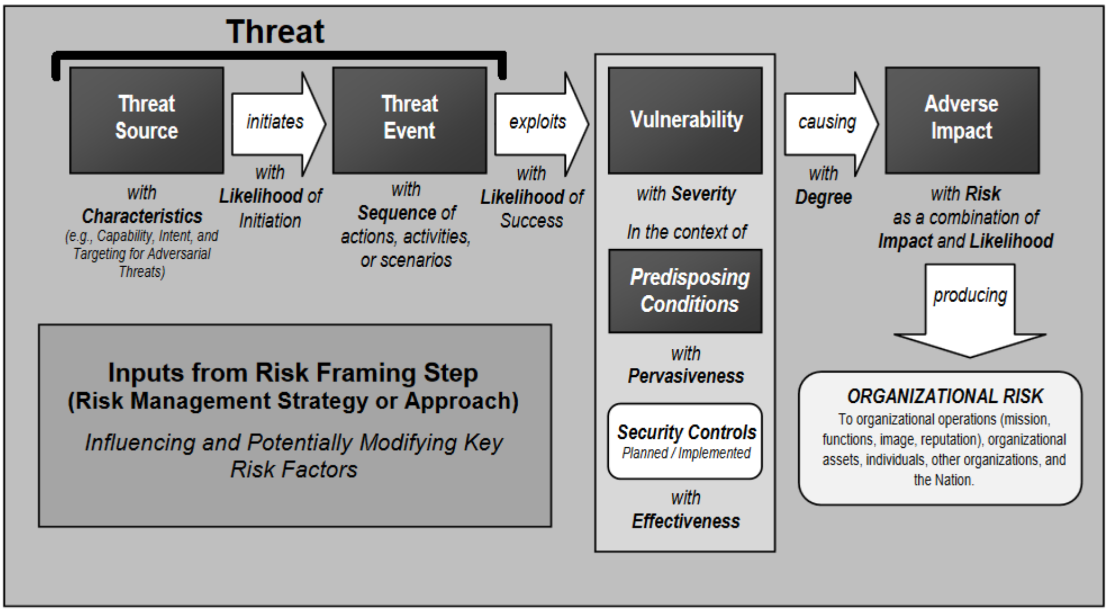

Threat
Threat Source
initiates
with Likelihood of Initiation
Threat Event
exploits
with Likelihood of Success
Vulnerability
causing
with Degree
Adverse Impact
with Characteristics (e.g., Capability, Intent, and Targeting for Adversarial Threats)
with Sequence of actions, activities, or scenarios
with Severity
In the context of
Predisposing Conditions
with Pervasiveness
Security Controls Planned / Implemented
with Effectiveness
with Risk as a combination of Impact and Likelihood
producing
ORGANIZATIONAL RISK
To organizational operations (mission, functions, image, reputation), organizational assets, individuals, other organizations, and the Nation.
Inputs from Risk Framing Step (Risk Management Strategy or Approach)
Influencing and Potentially Modifying Key Risk Factors

Figure 2.5: Example of a generic risk model referencing diverse risk factors and their interactions [2].

After introducing the fundamental concepts on which risk assessment relies, it is finally time to describe the process as defined by the standard briefly. As shown in Figure 2.6 the risk assessment process comprises four ordered steps: preparation, execution, communication, and upkeep. The following paragraphs will introduce and define what actions those steps consist of.

Preparation Before beginning the execution of the risk assessment process, it is mandatory to have a preparation phase; such phase has the objective of contextualizing the execution to the organization's needs, environment, and model. A successful contextualization starts from the results produced by the risk framing phase thoroughly described previously. With the help of such information it is necessary to identify:

- The purpose of the assessment, which must be specified with a great level of detail so that the assessment can output sufficient information, thus supporting the eventual decisions for which the assessment was requested;

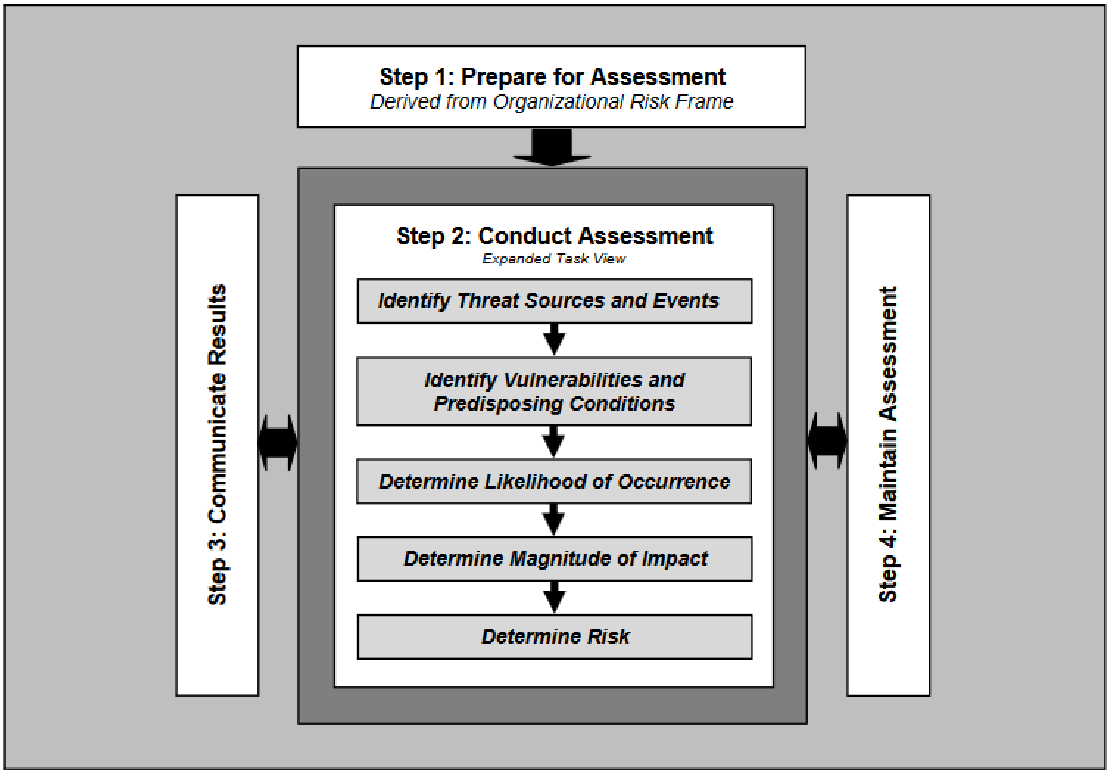

Step 1: Prepare for Assessment
Derived from Organizational Risk Frame
Step 2: Conduct Assessment
Expanded Task View
Identify Threat Sources and Events
Identify Vulnerabilities and Predisposing Conditions
Determine Likelihood of Occurrence
Determine Magnitude of Impact
Determine Risk
Step 3: Communicate Results
Step 4: Maintain Assessment

Figure 2.6: Schema of the risk assessment process and its phases [2].

- The scope of the assessment, both along the time dimension, hence the time frame to which the analysis is bound too, and the spatial dimension, which refers to the components of the organization (e.g. tiers, components of the architecture, specific security controls) to be analysed;

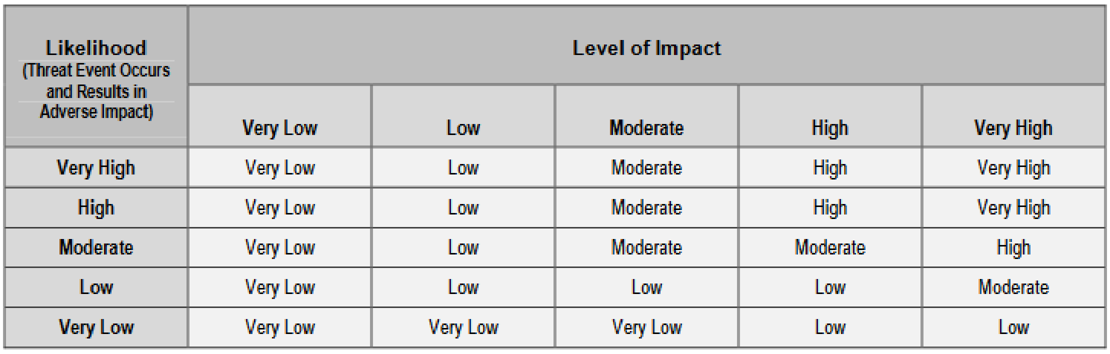
- The risk model and the analytic approaches to be used during assessment; while the model defines the risk factors to be assessed (as described in the previous section and as shown in Figure 2.5, the analytic approach defines both a way to assess information (i.e., quantitative, qualitative, semi-quantitative) and a perspective from which conducting the assessment (i.e., threat-oriented, asset/impact-oriented, vulnerability-oriented);

- The information sources that will be used as inputs to the assessment process; this data should be as comprehensive as possible, providing the necessary coverage for every risk factor integrated with the risk model (e.g. threats, vulnerabilities, likelihood, impacts, etc.);

- The assumptions and constraints associated with the assessment further explain the information contained in the risk model, contextualizing decisions and increasing the reproducibility and repeatability of the assessment's results.

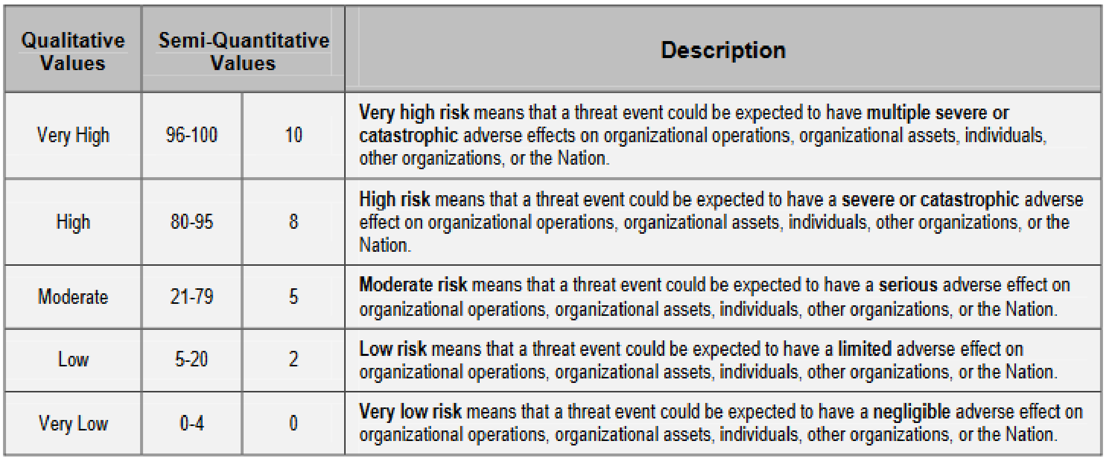
Execution The second step of risk assessment consists in conducting the assessment itself. This procedure aims at producing a list of information security risks that can be prioritized by risk level and used to inform risk response decisions. The organization is hence required to analyze threats, vulnerabilities, impacts, and likelihoods extracted from the information sources defined in the preparation step. The execution step, in fact, also implies gathering information from such sources expecting an adequate coverage of the factors taken into consideration. Going further into detail, conducting the assessment includes the following specific tasks:

- Identifying threat sources relevant to the organization, that is, starting from the gathered information, to define which threats are targeting or may be targeting the enterprise;
- Identifying threat events that may be produced by those sources, compatibly with the threat's characteristics (i.e. capabilities, intents, targeting);
- Identifying vulnerabilities affecting the organization that may be exploited by the identified threat sources via the possible threat events while defining the predisposing conditions that facilitate or impede successful exploitation;
- Estimate the likelihood of initiation of the identified events as well as the likelihood of success of such events;
- Estimate the impact on the organization's assets and operation in case of success of the threat events;
- Estimate information security risk as a combination of likelihood and impact of exploitation, specifying all the uncertainties associated with the estimation.



Likelihood (Threat Event Occurs and Results in Adverse Impact)	**Level of Impact**				
	Very Low	**Low**	**Moderate**	**High**	**Very High**
Very High	Very Low	Low	Moderate	High	Very High
High	Very Low	Low	Moderate	High	Very High
Moderate	Very Low	Low	Moderate	Moderate	High
Low	Very Low	Low	Low	Low	Moderate
Very Low	Very Low	Very Low	Very Low	Low	Low

Table 2.1: Example of combining qualitative assessments of likelihood and impact to produce a qualitative risk value [2].



Qualitative Values	**Semi-Quantitative Values**		**Description**
Very High	96-100	10	**Very high risk** means that a threat event could be expected to have **multiple severe or catastrophic** adverse effects on organizational operations, organizational assets, individuals, other organizations, or the Nation.
High	80-95	8	**High risk** means that a threat event could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.
Moderate	21-79	5	**Moderate risk** means that a threat event could be expected to have a **serious** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.
Low	5-20	2	**Low risk** means that a threat event could be expected to have a **limited** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.
Very Low	0-4	0	**Very low risk** means that a threat event could be expected to have a **negligible** adverse effect on organizational operations, organizational assets, individuals, other organizations, or the Nation.

Table 2.2: Conversion of risk values between different analytic approaches [2].

Communication The communication step involves sharing risk-related information obtained from the assessment. By correctly communicating those results, decision-makers all across the organization obtain updated knowledge regarding the organization's risk; as a consequence of this, those entities have the preparation needed for taking risk-related decisions. Depending on the organization's policies the method of communicating such information can take the form of executive briefings, risk assessment reports, or dashboards.

Upkeep Finally, the fourth step of assessment is to maintain the assessment. Maintaining the risk assessment means continuously incorporating any changes detected via risk monitoring, the process that allows determining the effectiveness of the risk response strategies implemented, as well as detecting the changes to information systems and their environment that impact risk. In practice this phase of the process requires monitoring the risk factors periodically, understanding why and how those factors change over time and consequently updating the risk assessment components related to monitoring and integrating the newly obtained knowledge.

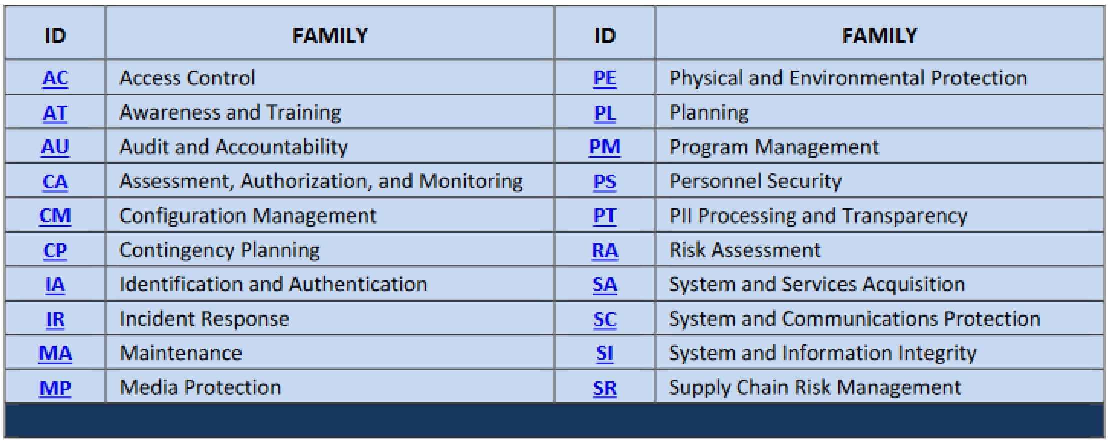
NIST SP 800-53

Titled “Security and Privacy Controls for Information Systems and Organizations” the standard [3] serves the purpose of a catalog of security and privacy controls for information systems and organizations. Those security controls aim to protect the organizational operation assets and everything included in the organization environment from an extended set of risks and threats.

The general knowledge of those controls is supposed to be helpful during the risk assessment process since allows for a better understanding of the security architecture of an organization as well as an easier detection of lacking security measures that therefore may increase risk.

Apart from the categories listed in Table 2.3 it is also useful to specify the different approaches for the control implementation, security controls can in fact be: common (inheritable), system specific, or hybrid.

Common controls are those whose implementation produces a capability that can be beneficial to multiple systems or programs via inheritance. In other words, an inheritable control will provide protection to a system despite being developed, implemented, and maintained by entities that are not directly responsible for said system. An example of this type of control is can be the security personnel, whose physical protection is inherited by information systems for instance. Nevertheless, it is extremely important to underline how this kind of control may constitute a single point of failure, so an excessive reliance on such controls should be considered carefully.



ID	FAMILY	ID	FAMILY
AC	Access Control	PE	Physical and Environmental Protection
AT	Awareness and Training	PL	Planning
AU	Audit and Accountability	PM	Program Management
CA	Assessment, Authorization, and Monitoring	PS	Personnel Security
CM	Configuration Management	PT	PII Processing and Transparency
CP	Contingency Planning	RA	Risk Assessment
IA	Identification and Authentication	SA	System and Services Acquisition
IR	Incident Response	SC	System and Communications Protection
MA	Maintenance	SI	System and Information Integrity
MP	Media Protection	SR	Supply Chain Risk Management

Table 2.3: Security and privacy control families as categorized by the standard [3].

Controls that are not implemented this way are either system specific or hybrid; the first ones being a primary responsibility of the system owner, those control may introduce risk in the architecture when the implemented control is not interoperable with the common controls present. Hybrid controls on the other hand are a middle ground between the two typologies: one part of the control entity is common and, consequently, inheritable, while the other is system specific; hence the risk deriving from the incompatibility among system-specific controls and common controls is removed since the two components are now made compatible by definition.

An evident conclusion from the standard is that security controls play a fundamental role inside the security architecture of an enterprise, thus heavily influencing risk factors considered during the risk assessment process affecting measures of likelihood and impact, and even introducing or removing new risks to the system with their absence or their presence. Hence, the analysis of security controls is mandatory in a risk assessment process.

2.2.2 ISO/IEC 27000 Series

The ISO/IEC 27000 series is a collection of documents created by the ISO and published by both ISO and IEC. The series is titled “Information Security Management Systems (ISMS) Family of Standard” and its purpose is to group international regulations aimed at protecting information gathered, stored, and processed by organizations. By implementing the practices and rules, enterprises can develop their own systems and processes for data protection against cyber-attacks, human errors, or any other vulnerability that may affect the organization and its infrastructure. Further in detail, the documents of the series: define requirements for the creation of systems dedicated to information processing and its security as well as requirements for the entities that certify those systems; provide guidelines for designing, deploying, maintaining, and upgrading such systems; provide instructions about the correct usages of different information systems.

One of the most important aspects of this series of publications is its versatility, they are, in fact, applicable to organizations of every category and size; the existence of such standards is a great incentive for the evaluation of risk related to information systems since it provides companies with an exhaustive material and suggestions on how to handle risk depending on the needs of the organization.

ISO/IEC 27001

Since an extensive analysis of all the standards of the series would require an excessive amount of time and cover specific topics that are out of the scope of this thesis, the research focuses on the ISO/IEC 27001 document, which is titled “Information technology - Security techniques - Information security management systems - Requirements”, and has as main focus the definition of the requirements of the information systems that aims at compliance with the ISO/IEC 27000 family of standards.

Information is considered a valuable asset for the organization, it is a fact that many modern organizations rely on data as the main way to profit; with this premise, it is necessary also to state that this information needs an infrastructure that allows for gathering, processing, and storage of it. Information systems represent such infrastructure but come with the cost of introducing risk for the organization, especially in a dynamically evolving context as the cyber world. The standard in analysis has as its objective the protection of the data by means of the so-called CIA Triad: to protect the information it is mandatory to guarantee its Confidentiality, its Integrity, and its Availability, hence information systems are responsible for managing such data must provide those features.

The aspect that mostly characterizes this standard and links it to the subject of risk management and analysis is the complete coherence with the risk management approach defined in the documents described in Subsection 2.2.1, specifically during the design of systems, a risk assessment process is required for identifying and evaluating risks with the subsequent definition of a Statement of Applicability for the system under exam. The Statement of Applicability is a document that basically serves the purpose of listing the security controls deemed to be necessary with motivations for their inclusions, those controls are taken from another part of the standard, “Annex A”. While “Annex A” could not be included in the thesis since protected by copyright laws, it is possible to say that it is a table containing all the categorized security controls that systems in compliance with this family of standards must implement. According to the organization’s mission and to the information system’s purposes many controls may not be applicable or needed, in that case, the Statement of Applicability must contain explanations for those decisions.

2.2.3 Digital Operational Resilience Act (DORA)

The Digital Operational Resilience Act is a legislative proposal published by the European Commission that contains a package of measures aiming at regulating Digital Operational Resilience in the European financial services sector [4]. The proposal consists of a series of legislative acts with the purpose of further supporting and enabling the potential of contemporary digital finance, assisting innovation and competition while at the same time diminishing the risk bound to the technology by providing well-suited safeguards against cyber attacks and other threats. Another reason behind those proceedings is the safeguard of clarity and coherence (especially with the interactions with other European directives) by introducing a single rulebook for the sector. The document targets a broad range of financial institutions, adding new categories of companies that have not been subject to regulations in the ICT field in the past such as cryptoasset service providers.

DORA is under scrutiny by the European Parliament and as of May 2022 a provisional agreement has been reached by the European Council and the European Parliament [5], once formally adopted it is expected to become operational by 2024.

In the following sections, the main provisions of the legislation are summarized.

Scope of Application (art. 2)

As stated previously, the legislation aims to cover all the institutions, working in the financial sector, that rely on ICT for their business, the so-called "Fintech" enterprises. In the last few decades technology has gained a pivotal role in finance hence determining a critical relevance in the daily functions of those entities. Digitalisation covers, for instance, payments, which have increasingly moved from cash and paper-based methods to the use of digital solutions. Other examples of this paradigm shift can be found in electronic and algorithmic trading, peer-to-peer finance, credit rating, and insurance underwriting. Finance has not only become largely digital throughout the whole sector, but digitalisation has also deepened interconnections and especially dependencies within the financial sector and with third-party infrastructure and service providers [4]. The presence of these interconnections and dependencies between financial entities, third parties, and ICT system has been identified as a potential vulnerability that in case of cyber incidents to any of the entities taken under analysis, could facilitate its spread to the entire system [6]. Therefore, significant breaches do not only affect the financial entities isolated but propagate across the financial transmission channels and could potentially produce negative consequences to the entirety of the European financial system, hence the need for comprehensive legislation from the European Union itself.

Roles (art. 4)

According to DORA the enterprise's governance must provide guidance for the risk management framework as well as ensure vigilance and effectiveness of the principles of cybersecurity. On the other hand, management has full responsibility and accountability over risk management implementation; the specific tasks for which management is responsible are:

- Clearly assigning roles and responsibilities for the ICT related functions;
- Periodically monitor, review and update the risk management strategy of the financial entity, properly adjusting risk factors and tolerances, as well as train on a regular basis to keep up to date skills and knowledge related to ICT risk;
- Ensure resource allocation for investments related to ICT, including expenses for training.

Risk Management (art. 5-14)

The legislation requires financial entities to develop processes targeting their ICT systems to guarantee resilience. Management should assign key roles to fulfill the following tasks:

- Proper identification, classification, and description of all the business functions relying on ICT, the assets related to said functions, and the sources of risk. Risk assessment is required to be performed annually;
- Proper implementation of technical mitigations to prevent risk;
- Detection of possible points of failure and suspicious activities by means of periodic testing;
- Development, testing, and maintenance of ICT Business Policies and ICT Disaster Recovery plans for response and recovery from disruptive incidents;
- Responsibly disclose incidents to partners, clients, and eventually the public and study causes of incidents as well as the effectiveness of the protection measures implemented.

Incident Handling (art. 15-20)

Financial entities are required to establish and implement an ICT related incident management process that involves detection, management, and notification of the incidents. While all the incidents should be classified and assessed, only major incidents should be reported to higher levels of management and to the competent authority. DORA specifies timelines and levels of details for reporting those incidents to said authorities. Firstly, an initial notification must be sent by the end of the business day; no later than a week after the initial notification, an intermediate report of the incident must be produced, shared, and complemented by notifications for every significant update obtained after said report; finally, no later than a month after the previous report, a final, complete report containing the analysis of the incident's cause and actual impacts, must be delivered to the authorities.

ICT Systems Testing (art. 21-24)

DORA defines general requirements for testing the performances of digital operational resilience with the purpose of assessing the company's preparation towards ICT related incidents. Testing should be intended as an integral part of the risk management framework of the financial entity and should be performed following a risk-based approach, hence taking into account significant risk factors during those procedures. All the critical ICT systems should be tested at least yearly and the testing programme should include a full range of different tests, including but not limited to: vulnerability assessments and scans, network security assessments, physical security reviews, source code reviews where feasible, compatibility testing, performance testing and penetration testing. Particularly, threat-led penetration testing (TLPT), where testers carry out tactics, techniques, and procedures used by real-world threat actors against systems, should be performed on critical functions and services (including third-party provided services) deployed on live production systems at least every three years. The results of such testing should then be provided to the competent authority but also serve as guidance for applying effective risk management controls to reduce impacts and consequently risk.

ICT Third Party Risk (art. 25-39)

Regarding risk profiles derived from ICT services suppliers, DORA aims at guaranteeing accurate monitoring of risk, establishing key elements of the collaboration with third parties. Contracts with such entities must contain: a complete description of the services provided; details of the data processing centers; descriptions of the guaranteed SLA accompanied by performance objectives; dispositions for data confidentiality, integrity, and availability; rights for inspections of the systems from the financial entity or authorized third parties; clear rights for contract resolution and dedicated exit strategies.

2.2.4 The TIBER-EU Framework

Although not directly providing guidelines on risk management, the TIBER-EU Framework [7] defines the foundations upon which the DORA is built, as well as provides interesting overviews of different threat intelligence-based approaches for system testing that aim at improving system resilience. For these reasons, and for the sake of completeness, a brief description of this structure is therefore provided.

Conjointly developed by the ECB and EU Central National Banks, the framework was published in 2018. The purpose of the work is to provide to regulatory authorities and European entities (especially those involved in the EU financial infrastructure) guidelines on how to interact and cooperate with threat intelligence providers and red teaming[1] services providers focusing on improving the resilience of information systems against cyber attacks.

The core concept of the approach described by the TIBER-EU framework is conducting testing of the entities' critical live production systems in an intelligence-based fashion, thus, by mimicking threats with their techniques, tactics, and procedures (TTPs). By following the threats' approaches, the systems are targeted and attacked with real-world adversarial practices, allowing the assessment and evaluation of the effectiveness of protections, detection mechanisms, and countermeasures implemented on said systems.

[1] https://en.wikipedia.org/wiki/Red_team

Without going into particular details, testing compliant with the TIBER-EU structure will therefore require a specific separation of roles and responsibilities which can be summarized as follows:

- *Threat Intelligence provider*, the organization responsible for performing reconnaissance on the entity with the purpose of identifying and modeling the threats actors that could target said entity;
- *Blue Team*, the people within the entity that undergo the process of testing, together with the information systems. The process will test their capabilities in preventing, detecting, and responding to adversarial actions and generally starts without their knowledge, nevertheless, some forms of testing (e.g. purple teaming) may involve the acknowledgment of the team;
- *Red Team*, the company that provides red teaming services, that is, to say, carries out the attacks against the entity's infrastructure following the TTPs of the threats that have been identified by the threat intelligence provider;
- *White Team*, a restricted number of people inside the targeted entity that are aware of the testing and are partially responsible for the management of the test;
- *TIBER Cyber Team*, an external team from the regulatory authority that is responsible for monitoring the testing process, guaranteeing that it follows the framework's requirements.

The framework's requirement for threat intelligence dictates a particular focus on the need for methodologies that allow the gathering of such information, as well as definitively establishing another important category of information (i.e. threat information) that should be taken into account when aiming at securing information systems. Consequently, it seems reasonable to assume that this kind of knowledge could also provide significant insights from a risk assessment perspective.

One last consideration about the TIBER-EU approach must be taken from the perspective of risk management; introducing a threat lead testing on live production systems could in fact bring disruption to the business functions of the entity since events such as unexpected crashes, denial of services or damage to the systems could happen despite the precautions deployed. Therefore before conducting any testing that involves threat-based red teaming activities it is of considerable importance to perform once again a risk assessment of the process. By doing so the derived risk will be highlighted during the entirety of the process, allowing the target entity to take the right precautions in order to respond to unexpected events with a strategy that minimizes the consequences of such events, and, therefore, risk.

Chapter 3

Problem statement

In the context of modern enterprises and organizations, information is available in great quantities. Ranging from internal reviews and investment strategies to information system logs and security controls, many processes can be considered as sources of data that could be meaningful from a cyber risk management perspective; however, such a wide amount of knowledge is intrinsically challenging for human beings to analyze aiming for the discernment of information that is actually relevant for framing, assessing and analyzing risk. To accomplish this task, establishing automated processes executed by information systems is deemed a reasonable solution. With digital systems assisting or even substituting humans in the information collection and refinement process, both efficiencies, as well as accuracy can be improved, leading to better results and more clarity in the risk management process.

3.1 Assumptions

The work contained in this thesis relies on several basic assumptions starting with the fact that introducing automation in the risk management framework can bring significant improvement to the entire process pipeline, reducing time and resources, specifically human resources, dedicated to information-related tasks, as well as increasing the level of performance reached by such activities. Given these premises, it is also reasonably assumed that companies and organizations invested in cyber risk analysis procedures, such as ensuring firms or risk management services providers, but also less specialized enterprises that still require an internal risk management pipeline, may find great interest and benefits in automation of such processes.

Other assumptions have been dictated by the constraints implied by the development context of this thesis (explained in Section 3.2): given the external approach chosen, it is assumed that enough public information is available for each target when executing data collection.

Lastly, the documents studied and presented in Section 2.2 constitute an integral part of this thesis's research and development approach. Therefore it is taken for granted that the methodologies therein described actually represent real-world best practices validated with enterprise experience.

3.2 Constraints and Limitations

The research was conducted without the direct involvement of third parties such as organizations and enterprises that could provide access to sources of information and validation of data specifically constructed for risk management purposes such as assessment and analysis; therefore, any result obtained must rely exclusively on publicly available information. As a consequence of this limitation, the perspective of the research has been adapted to that of an external attacker with no information outside the openly available and no access to the internal infrastructure and systems of the target. Notably, this approach shares similarities with a vulnerability assessment procedure when the scope is limited to publicly accessible assets, thus enabling a wide range of techniques; the most prominently used in the context of this thesis being vulnerability scanning with the tools described in Subsection 5.2.2.

Other limitations can be identified in the context of information enrichment by means of machine learning as introduced in Subsection 5.3.1 and further explained in Chapter 5; a crucial step for many machine learning processes is the availability of proper datasets for the purpose of model training. However, as thoroughly studied by Cremer et al. [8], datasets related to cyber risk constitute a smaller portion of cybersecurity data collections, which, in turn, are too little or no availability from publicly available sources.

One last consideration regarding the tools used during the development of this thesis, which is described in Section 5.2, is that every piece of software mentioned is either open source, hence freely available for download, or has been employed in its version for free evaluation purposes which generally implies a limitation on the available features. However, it was still possible to obtain more general yet still valuable information even from the incomplete versions.

3.3 Research Goals

Given the aforementioned premises, this research aims to explore state-of-the-art technologies for gathering and processing risk-related information, aiming for better integration of automatic techniques. Another prerequisite of this objective is identifying sources of information that can be linked to cyber risk. Therefore, part of the research is dedicated to the discovery and refinement of specific data sources that are able to provide the required information.

The culmination of the exploration of techniques and the research of proper sources of information is the integration of these two portions into an automated framework. This software framework is able to gather information from the discovered sources by employing the methodologies studied and following a predictable logic, which is defined by a set of rules. As a result of this union of concepts, a new tool has been developed, allowing for a reasoning-driven collection of information that not only is able to gather new data based on previously obtained information but is also capable of producing novel, enriching knowledge without the need for user interaction.

Chapter 4

Analysing and Modelling Cybersecurity Standards

Studying the standards to which many contemporary enterprises and organizations of several different sizes comply, is an excellent starting point for extrapolating important key business concepts relevant to a risk assessment context. Assuming that these documents are optimal, having an insight into the best risk prevention and mitigation practices across all the different compartments of a business allows for identifying what is needed to maintain a high level of protection and conduct business operations safely. Therefore, the main objective of this Chapter is to recognize various important concepts that should be taken into account and evaluated, when possible, when conducting a risk assessment; once defined, this information can be labeled depending on its characteristics in order to facilitate the process of transforming it into more direct methods of evaluation (i.e. business surveys, tools for direct analysis) that can eventually be automated.

4.1 Classification of information from standards

During this phase, a range of common properties of the studied information has been identified to partially model the data, thus introducing a clear classification of such knowledge based on its different specifics.

Additionally, this modeling approach of introducing common and generic meta-information is well suited for integration with other reasoning approaches that may take advantage of it in order to execute different actions that depend on the context of information. For example, information from different sources may have different reliability inside the system, triggering different paths of action.

A simple labeling model has been introduced to achieve this classification. Every concept is therefore associated with one (or less frequently, more than one) tag for the different information properties. These properties are:

- *Source*: the property that defines where the information is expected to be obtained. The possible labels are *company, trusted third parties, authorities, penetration testing, publicly available.*
- Purpose: the property that defines what the information is used for. The possible labels are *quantification, evaluation (i.e. assessing company parameters from an external point of view), likelihood estimation, impact estimation.*
- *Confidentiality*: the property that specifies which level of secrecy is expected to be associated with the information. The possible labels are *confidential, restricted, internal, public.*
- Impact: the property that specifies what kind of impact the information is related to. The possible labels are *economical, business process, image, sanctions, direct, indirect.*
- *Acquisition Phase*: the property that indicates which phase of the risk management process is expected to produce the information. The possible labels are *framing, assessment, response, monitoring.*

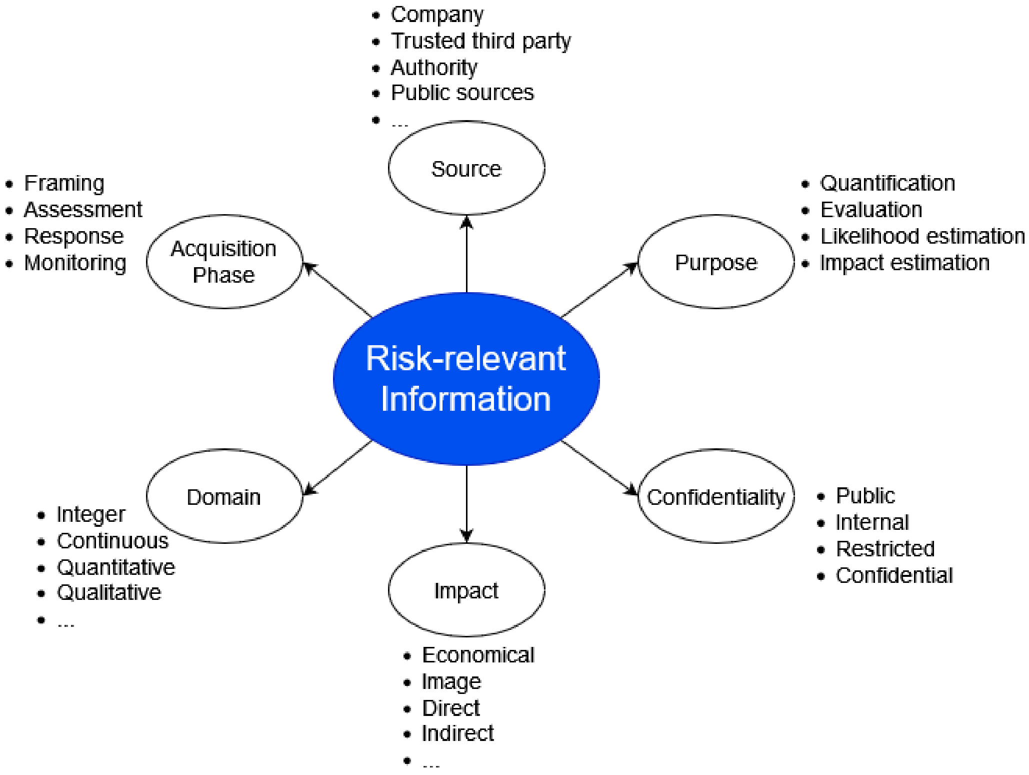


Figure 4.1: High-level representation of how the information is modeled with its main properties.

- *Domain*: the property that defines the format in which the information could be represented. The possible labels are *integer, continuous, semi-quantitative (i.e. ranges such as 0-100, 0-10, etc.), qualitative (i.e. non-numeric ranges such as "very low to very high").*

4.2 The concepts

Another important aspect of the main risk-related concepts that have been identified is that they fall under one of five macro-areas of risk management. This further separation is extremely useful since it allows to split off a particularly vast scope of analysis into smaller areas that can be processed separately in the majority of cases, enabling the modularization of the information-gathering processes. The various concepts identified for each subtopic have been presented in the following subsections.

4.2.1 Assets

One of the first areas of interest for every risk management process is the identification of assets; this operation poses the basis for numerous other activities related to the assets themselves: from vulnerability assessments to mitigation planning, having an accurate representation of the parts of the company system and what constitutes them is fundamental in order to keep the deriving cyber risk under control. For these reasons, it has been chosen to identify possible concepts from which to extract risk-relevant measures related to the assets introduced in the next paragraphs.

Volume of cyber-related assets

An indicator, possibly proportional to the volume of all the assets, is the magnitude of resources related to information systems. Examples of cyber-related assets are devices connected to the network, used for processing and/or storing data, and data itself. Examples of non-cyber-related assets are physical devices lacking any particular digital control system that may be compromised by external entities.

Labels: *assessment, 0-100, internal/restricted, company/trusted third-parties, quantification/impact estimation.*

Volume of exposed cyber-related assets

An indicator, possibly proportional to the volume of the cyber assets, of the magnitude of cyber-related assets that are accessible from external entities. Assets of this type are, for example, web servers hosting the company's websites and provided services.

Labels: *assessment, 0-100, public, public sources, likelihood estimation*

Volume of not exposed cyber-related assets that are accessible internally

Similarly to the previous one, this measure constitutes a proportional indicator of the volume of the assets that can be accessed by internal, generic, and low-privilege employees but should not be accessible from external entities. Examples of this type of asset are company network shares for data retrieval and internal proprietary software/technologies.

Labels: *assessment, 0-100, internal, company, quantification/likelihood estimation*

4.2.2 Vulnerabilities

Another important area of interest that is considered during the risk management process, especially during the assessment, is identifying, evaluating, and prioritizing vulnerabilities affecting the information systems that comprise the business architecture. Indeed, it is easy to see that this phase is generally dependent on a previous asset identification process, both for scoping an eventual vulnerability assessment as well as providing priorities by means of the asset's importance to the discovered vulnerabilities. The following paragraphs describe possible concepts to be measured in the area of vulnerabilities.

Likelihood of misconfigurations

An indicator of how prone the company's information systems are to human errors and misconfigurations that lead to attacks. Although not completely, this measure could be derived from historical forensics evidence of attacks, calculating a proportion of the number of attacks due to human errors and the total number of attacks.

Labels: *monitoring, internal/restricted, company, likelihood estimation, 0-100*

Average exploitability of vulnerabilities

A measure that gives an insight into the ease of exploitability of the vulnerabilities present in the company's information systems. Several scoring systems have been developed for this purpose, the main one being the EPSS. The average could also be weighted on the value of the assets afflicted, and the scope of analysis could be sized depending on the necessary granularity.

Labels: *assessment, restricted, penetration testing/company, likelihood estimation, 0-1/qualitative*

Average impact of vulnerabilities

Strictly correlated with the relevance of the affected assets, this measure is aimed at estimating the impact that the discovered vulnerabilities may produce if exploited. Once again, for aggregating measures, the scope of the aggregation could be sized accordingly to the required granularity.

Labels: *assessment, company, restricted, impact estimation, 0-100/qualitative*

Average exposure of vulnerabilities

This indicator refers to the proportion of vulnerable systems (with respect to the number of vulnerable information systems, and, eventually the total number of information systems) that are accessible from external entities. The proportion could be weighted based on the vulnerabilities' impact and/or exploitability.

Labels: *assessment, restricted, penetration testing/company, likelihood estimation, 0-100*

4.2.3 Threats

Threat analysis is an area mainly independent of the others presented in these paragraphs. Investigating and collecting intelligence about the adversarial entities that may target the company or organization is an extremely important step during risk management, and in particular during risk assessment, because it allows contextualizing in a more precise way how and if risk sources will be exploited for adversarial action, thus providing context-aware estimates of likelihood and impact of threat events by defining intentions and capabilities of such threats.

The concepts presented in the following paragraphs describe measures and indicators that may be used to evaluate threats in a risk management context and consider both internal and external threats. Apart from the standards presented in Chapter 2, many interesting concepts used for modeling threats have been derived from a report produced by Sandia National Laboratories titled “Cyber Threat Metrics” [9].

Type of threats

A general indicator of the class of threat actors that may target the organization, weighted aggregation (average) can be performed by establishing a numerical scale and weighting based on the threat’s commitment to initiate threat events (i.e. the frequency of attacks).

Labels: *framing, trusted third parties/authorities, restricted/public, evaluation, qualitative*

Frequency of threat events per threat

Based on historical evidence of attacks, this indicator measures the commitment of a threat actor in targeting the organization.

Labels: *monitoring, company/authorities, restricted/public, likelihood estimation, numerical*

Capabilities of threat actors

A qualitative indicator of the resources (i.e. economic, infrastructural, and technical) that a threat actor is willing to employ when targeting the organization. The granularity can be kept for singular threat actors or aggregated to produce average results.

Labels: *framing, trusted third parties/authorities, restricted/public, likelihood estimation/impact estimation, qualitative*

Employees degree of freedom on systems

A measure aiming at evaluating the impact derived from internal threat actors (i.e. employees) that are targeting the organization's assets. This indicator is also useful for measuring the likelihood and possible impacts of unintended human errors produced by the company's employees.

Labels: *assessment, company, internal, likelihood estimation/impact estimation, qualitative*

4.2.4 Mitigations

An appropriate risk management strategy always includes implementing security controls and policies that reduce the company's exposure to risk. Evaluating, maintaining, and eventually improving these security measures over time is crucial for a structured and performing risk reduction. The starting point of this process is obviously the evaluation of such mitigations, which is usually performed after their implementation, thus during the monitoring phase. From the risk assessment perspective, having an insight into the effectiveness of the security measures implemented could be particularly useful for estimating risk factors such as likelihood and impact. Therefore, the following paragraphs present concepts related to evaluating the mitigation strategies implemented.

Average patch response time

An indicator of the efficiency of the company for what concerns the remediation of newly discovered vulnerability. A low average patch response time is key to reducing the window of opportunity for threat actors to initiate threat events.

Labels: *response, company, internal, evaluation, numerical*

Mean time to detect

An indicator of the efficiency of the company for what concerns the detection of ongoing attacks. It is crucial to detect any ongoing adversarial action in the lowest timeframe possible to employ remediation strategies and reduce the impact of the attack.

Labels: *response, company, internal, evaluation/impact estimation, numerical*

Incident rate trend

Based on historical attack data, identifying trends in the number of successful threat events allows for evaluating the performances of the implemented mitigations: a decreasing trend could represent more effective security measures. In contrast, an increasing trend could mean the opposite.

Labels: *monitoring, company, internal, evaluation, qualitative*

Workforce dedicated to mitigation and incident response

The proportion of staff that is employed for purposes of mitigation, patching, and incident response. The proportion can be calculated over the number of total employees as well as the number of IT-specialized employees.

Labels: *framing, company, internal, evaluation, 0-100*

Company budget dedicated to mitigation and incident response

The proportion of economic resources invested for the purposes of mitigating cyber risk by means of security policies and controls, response strategies, and vulnerability patching.

Labels: *framing, company, internal, evaluation, 0-100*

4.2.5 Staff Awareness

The last area of interest for risk-relevant concepts analyzed in this chapter covers the technical IT and cybersecurity competencies of the company's employees. This includes technical and low-level staff, and management personnel since employees at every level may introduce risk to the organization both intentionally and unintentionally. In particular, the exigence of evaluating the risk-awareness and security training of the people executing and managing business processes on a day-to-day basis comes from the fact that more and more incidents have been caused by the negligence and maliciousness of insiders costing hundreds of millions to organizations [10].

The awareness of the staff can be seen as a combination of two sub-factors: knowledge, being the understanding of the systems and the processes under a security and risk-aware perspective; and involvement, which defines how well the staff integrates and participates in the business processes by means of improvements and suggestions. Both those sub-factors need to be evaluated.

Employees average level of training for cybersecurity

A measure used for estimating the likelihood of accidental internal events that can produce negative consequences for the organization. This measure is based on the assumption that higher training in cybersecurity best practices diminishes the likelihood of employees being the cause of accidental adverse events connected to information systems.

Labels: *assessment, company, internal, likelihood estimation, qualitative*

Periodicity and duration of training

These measures are supposed to evaluate the quality of the training received by the company's staff. Performing training activities with high periodicity generally implies having employees with up-to-date knowledge about novel threats and techniques used by attackers, while having more prolonged periods of training is assumed to be a synonym for transmitting higher volumes of knowledge and a better understanding of it. Evaluating the quality of training is useful for estimating how likely employees are to commit errors that may negatively impact the company.

Labels: *framing, company, internal, likelihood estimation, numerical*

Dependency of management from technical employees

This concept has the purpose of evaluating the ability of the management portion of the employees to be independent of the low-level employees; in other words, it is an indicator of the leverage that generic workforce employees have on the management, hence on the company and its business operations. A company that is highly dependent on its low-level employees is automatically more vulnerable to insider threats. On the contrary, companies where a clear separation both in privileges and responsibilities between management and the workforce is present, may be less affected by employees turning into adversaries.

Labels: *framing, company, internal, impact estimation, qualitative*

4.3 Third parties and Supply Chain

A separate discussion has to be made for the entities that the target enterprise relies on. Nowadays, most organizations and companies of all sizes depend to some extent on other service providers. Being the logistics providers, cloud services providers, or even consulting services, to name a few, they all constitute a possible cyber risk for the main organization, both directly and indirectly.

For this reason, the best way to evaluate the cyber risk these entities introduce to the business architecture is to apply the same concepts identified in Section 4.2, when possible and constrained to the available information.

Additionally, when dealing with the cyber risk constituted by third parties is particularly relevant to model the relationship between the main organization and these entities by means of trust, trustworthiness, and criticality of cooperation, three concepts that allow for a better evaluation of both the likelihood of adverse events happening and the impact that said events could have on the main organization.

Trust and Trustworthiness

When analyzing the relationship with a business partner, these two concepts help in establishing the approach to be used in order to minimize the derived risk. Trust represents the organization's point of view inside the partnership and basically describes the acceptable level of disruption of the services provided by the business partner, together with the level of confidentiality needed for data and business processes that have to be shared with the third party. On the other hand, trustworthiness represents an evaluation, usually based on possible conflicts of interest and historical data such as incident records and previous business relationships, of the quality and reliability expected for the services provided by the partner. These measures may act as a weighting factor to the cyber risk assessed with the approach described previously.

Labels: *framing, company, internal, evaluation, qualitative*

Criticality of cooperation

Another important concept that needs to be evaluated is how reliant the main company is on these third parties. Quantifying the level of dependability from a partner allows for a better representation of the impact that the disruption of said services could have on the main organization. Factors influencing this concept are the flexibility of the business contract stipulated with the partner, the possibility of modifying the contract terms without a premature resolution, or, in case of premature resolution, the estimated losses deriving from this resolution, and the presence of competitors that are able to promptly replace the business partner in order to reduce disruption periods.

Labels: *framing, company, internal, impact estimation, qualitative*

Chapter 5

Building Risk Analysis Models from Publicly Available Information

Since its creation, the internet has covered a fundamental role in sourcing and spreading every kind of information, quickly becoming the most prominent source of knowledge for many aspects of everyday life. From historical records to technical know-how, an extremely vast amount of information categories is freely accessible online within seconds; among these categories, one arouses particular interest for the purposes of this work: information classified as OSINT. Open Source Intelligence includes a broad spectrum of different data sources; every company, organization, or institution of every size owns a certain online presence; it may be more or less evident depending on the purposes and missions of the entity but in a way or another interesting detail end up online. Ranging from simple websites and web services to law-mandated financial disclosure obligations, countless examples of data that could provide insights about an entity's exposure to cyber risk are available to anyone that knows where to look at.

Unfortunately, a huge volume of data does not necessarily imply improvements in a cyber risk assessment process. For starters, significantly increasing the amount of information to be processed and analyzed could drastically worsen the performance of the process itself, even if properly designed, streamlined, and automated. Another important factor is that not every piece of data available does actually bring significant insights about cyber risk exposure. For these reasons, one of the challenges that need to be addressed during open source information gathering is how to perform an educated pruning of data, or even more optimal, how to educate the collection process in such a way that only risk-related information is extracted.

The following Sections present an analysis of tools specialized in information gathering, detailing the type of knowledge they can extract, as well as different techniques employed for the automatic enhancement of this knowledge aiming at generating newly derived information to improve cyber risk assessment further.

5.1 Workflow

This research part followed a consistent workflow about the vulnerability assessment tools. After selecting a set of interesting and available tools, their features were tested against a common target. Considering that, in general, this kind of software follows a common "Scan then Report" approach, they are particularly well-suited for comparison.

Therefore, after completing the scanning routines of every software, a set of similarly structured documents describing the vulnerabilities discovered, in conjunction with additional details that differ from tool to tool. Once these sample reports from the vulnerable common target were obtained, they needed to be manually analyzed to understand which risk-related data they produce, how it is produced, and eventually how it was produced or how it could be extracted automatically from the given reports.

Following the identification of common grounds amidst the tools and especially significant data, the next step was to process the obtained information to enrich it with new data. To do so, the predominantly used technique was Machine Learning, particularly Natural Language Processing, as better described in Section 5.3.

The final output of the process then becomes a high-level definition of the

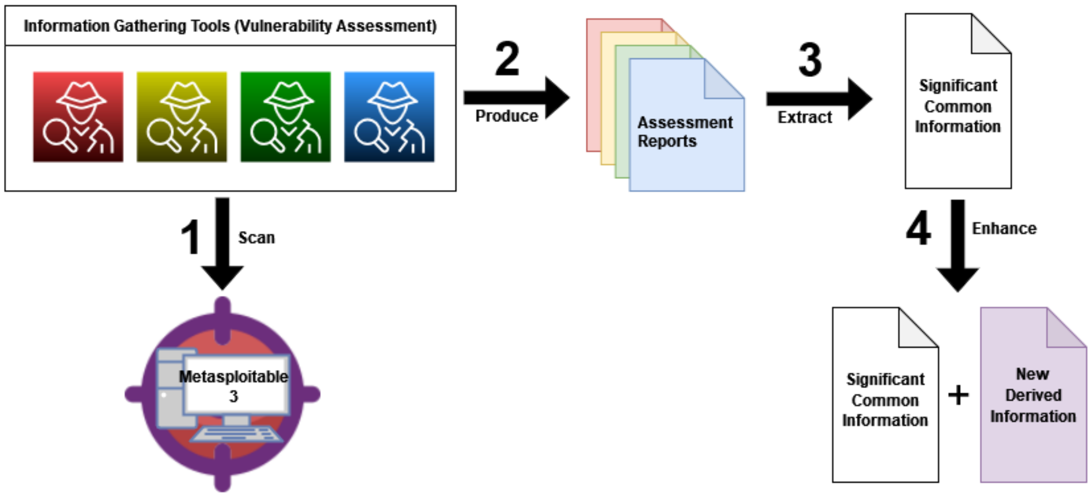

Information Gathering Tools (Vulnerability Assessment)
2
Produce
Assessment Reports
3
Extract
Significant Common Information
1
Scan
Metasploitable 3
4
Enhance
Significant Common Information
+
New Derived Information

Figure 5.1: High-level workflow for tool comparison, information extraction, and enrichment.

information obtainable from the vulnerability assessment tools, its structure, and how it was obtained, as well as eventual additional knowledge that can be produced from it after additional processing.

The reason behind this process is found in the techniques used during a risk assessment process and clearly related to the systems' vulnerability assessment of the organization under scrutiny. Since vulnerability assessment is a process usually conducted by third parties, it may be both time-consuming and costly. For this reason, an analysis of the data produced by it and the techniques (namely, the tools) employed, posing significant attention towards automation, could lead to major improvements in a task that, under certain aspects, could be considered pure information gathering.

5.2 Tools and Targets for OSINT through Vulnerability Assessment

5.2.1 Vulnerability Assessment

One evident deduction from the documents described in Section 2.2 is the necessity for the detection of vulnerabilities present in the organization's systems. Vulnerabilities are one of the key risk factors in the risk management framework's risk models. Hence, companies may find it extremely desirable to establish a process for weakness-related information gathering; the said process is, in fact, a Vulnerability Assessment.

Similarly to the risk assessment process, vulnerability assessments can be broken down into their individual phases:

- *Scope identification*, which is responsible for determining which portion of the organization's information systems will be tested and to which extent; this phase is highly correlated with asset identification, which is a necessary step in a risk management framework;
- *Prioritization* of the assets identified inside the scope of the vulnerability assessment, which consists in assigning a quantifiable value of importance to such resources, thus contextualizing them to the business environment;
- *Scanning*, the process of identifying the vulnerabilities affecting the entry points (e.g. the services present on or related to the assets in scope);
- *Reporting and mitigation*, the phase which processes the outputs of the scanning phase and based on priority determines, when possible, whether and how to reduce impacts and likelihood of exploitation of the identified weaknesses.

5.2.2 Vulnerability Assessment Tools

For the purpose of vulnerability assessment, a wide variety of tools and technologies, both open-source and proprietary, have been developed. Those technologies mostly aim at covering the scanning phase of the vulnerability assessment, with some exceptions that lightly cross the border to the other phases. The services such tools provide range from network mapping activities for partial asset and service discovery to fully-fledged vulnerability scanning, web applications, and operating systems assessment.

For the purpose of this thesis, a selection of the most popular [11], the industry-standard tool, has been tested to study the knowledge that such software could produce, highlighting differences and common factors. Following the list of the examined tools along with the characteristics identified.

Tenable Nessus

The Nessus tool[1] offers:

- Plugin-based vulnerability scans which provide for extensibility of features through the addition of new plugins, with each plugin aiming at covering: specific CVEs or specific CWEs;
- Remediation instructions, if possible, for identified vulnerabilities, with capabilities for aggregating multiple vulnerabilities under a single proposed solution;
- Proof of concept exploit for the vulnerability with examples outputs of successful exploitation on the target;

The output report in its detailed format contains, for each vulnerability of each target service (e.g., each scanned IP address, and each open port of that address): severity of vulnerability (CVSS[2] score); textual description of vulnerability; a qualitative risk factor (e.g. Low, Medium, High, or Critical); general remediation instructions when available (e.g. “Upgrade software to version x.x.x”); especially for HTTP services vulnerable endpoints with proof of exploit (request/response); ease of exploitability in the form of

[1] `https://www.tenable.com/products/nessus`

[2] `https://www.first.org/cvss/`

listing framework able to exploit the vulnerability (e.g. "Exploitable with Metasploit[3]").

Greenbone OpenVAS

Greenbone OpenVAS[4], similarly to the previous one offers:

- Plugin-based vulnerability scans which provide for extendability of features through the addition of new plugins, with each plugin aiming at covering: specific CVEs or specific CWEs;
- Remediation instructions, if possible, for identified vulnerabilities, with capabilities for aggregating multiple vulnerabilities under a single proposed solution;

The output report in its detailed format contains, for each vulnerability of each target service (e.g., each scanned IP address, and each open port of that address): severity of vulnerability (CVSS score); textual description of the vulnerability and qualitative evaluation of severity (e.g., Low, Medium or High); general remediation instructions when available (e.g. “Install the latest version”), as well as solution type (e.g., fixes provided by the vendor, mitigation, etc.); depending on the plugin the vulnerability report can contain verbose outputs of the detection that, in particular for HTTP services could contain vulnerable endpoints and actions performed on them; the descriptive impact of what consequences the exploitation of the vulnerability could lead to.

NMAP

As the tool name anticipates, Network MAPper[5] is a tool that falls into a different category with respect to the others; its scope of application is that

[3] `https://www.metasploit.com/`

[4] `https://www.openvas.org/`

[5] `https://nmap.org/`

of network scanning and discovery; thus, it specializes in the detection of open TCP/UDP ports on the target hosts. It also provides the extensibility of features through the execution of scripts aimed at fingerprinting such services. Those scripts could also allow for basic vulnerability detection that relies mostly on service versioning and fingerprinting. Given the versatility of the tool, it is also included in this list.

Arachni

The tool Arachni[6], developed mostly for web application assessment, targets exclusively the HTTP services exposed on common ports, detecting available paths and testing them for common vulnerabilities.

The output report in its detailed format contains a textual description of the vulnerability; a qualitative severity of the vulnerability (e.g. Low, Medium, High, or Informational); the confidence of the detection (e.g. trusted or untrusted); general remediation instructions; vulnerable endpoints with proof of exploit (request/response) with textual explanations of the performed actions and consequences.

Invicti Acunetix

The tool's[7] main purpose is the vulnerability assessment of HTTP services, and among its features it provides:

- Several output templates including reports structured for compliance to different standards such as NIST SP800-53 (Subsection 2.2.1) and ISO 27001 (Subsection 2.2.2);
- Software dependencies recognition and vulnerability assessment;
- Custom scripts integration.

[6] `https://www.arachni-scanner.com/`

[7] `https://www.acunetix.com/`

The output report generally contains a textual description of vulnerability; the descriptive impact of what consequences the exploitation of the vulnerability could lead to, when available; vulnerable endpoints with proof of exploit (request/response) with textual explanations of the performed actions and consequences; detailed remediation instructions and recommendations; severity of vulnerability (CVSS score) and qualitative evaluation of severity (e.g. Low, Medium or High); the confidence of the detection (e.g. verified if the software was able to exploit the vulnerability or another numerical confidence level).

Invicti Standard (ex Netsparker)

The tool[8] offers similar features as the previous one but with a much larger set of plugins and checks executed on the target.

The output report generally contains the same information as the previous tool with additional levels of detail, as well as additional descriptive evaluation of the ease of exploitation of the vulnerability.

5.2.3 Other Tools for Information Gathering

Information is the foundation stone for every process in the risk management framework; it is almost mandatory to establish methodologies for risk-related information retrieval that are as automated as possible for better efficiency, repeatability and especially for obtaining bigger volumes of data. In Section 5.2.1, many possible interesting sources have been defined with the introduction of vulnerability assessment tools; nevertheless, it is important to notice that such a category of instruments relies on procedures for asset identification that, although simple, have not yet been fully automated. For this reason, tools that take advantage of search engines and other forms of OSINT (Open Source Intelligence) need to be introduced as a starting point for the information-gathering workflow.

[8] `https://www.invicti.com/web-vulnerability-scanner/`

The most important tool used for this purpose is theHarvester[9], an open-source tool used for reconnaissance during penetration testing. The application of this software in the context of this thesis is its ability to obtain IP addresses, URLs as well as subdomains starting from as little information as a domain name by simply utilizing publicly available data. Once obtained these first entry points for analysis, the other tools can further increase the volume of risk-related information. However, the tool can also provide other types of knowledge such as email addresses, phone numbers, and names of people linked to the targeted organization that could offer starting blocks for different types of assessment not necessarily related to information systems vulnerabilities.

5.2.4 Targets

To perform the evaluation of the features provided by the different tools described in the previous section, a target of the vulnerability assessment must be specified. The selection of such a target has been performed with a series of prerequisites in mind:

- The target must be openly testable without violating any term of services or bringing disruption, damage, or any adverse impact to third parties;
- The target needs to be sufficiently complex in order to evaluate the highest number of features of each tool, thus providing greater insights over the obtainable information;
- The target must be the same for each tool under scrutiny, allowing to establish of a common ground for the comparison of the tools' outputs;
- The target must be available for the evaluation period in order to have a stable testing environment; in such a manner, the evaluation process could proceed without the need of repeating time-consuming vulnerability assessments failed for lack of availability of the target.

[9] `https://github.com/laramies/theHarvester`

These requirements led to the choice of a locally executed virtual machine hosting multiple vulnerable web services on different ports; specifically the system is a Ubuntu[10] based machine developed by rapid7[11] and called "Metasploitable 3"[12] [13].

After starting the machine with the software Oracle VirtualBox[14] and properly setting up the network interface, the machine was accessible on the local network, allowing for the scanning operations performed by the tools.

5.3 Enriching information with advanced techniques

As previously mentioned in Section 5.1 the workflow presented in this chapter included a step dedicated to the processing of information to enrich it, or, in other words, to produce novel valuable information for risk assessment purposes. As generating new meaningful and coherent data not trivially conductible, advanced techniques are required to be employed, aiming at tasks with relatively simple outputs such as categorization or regression.

5.3.1 Natural Language Processing

The main subject for exploration of information processing and especially information enrichment that has been performed was machine learning and its branch dedicated to natural language processing. After realizing that most of the information produced by the tools of Section 5.2.1 was in a textual format it has been decided to implement techniques that allowed to retrieve of additional data from such texts.

[10] `https://ubuntu.com/`

[11] `https://www.rapid7.com/`

[12] `https://github.com/rapid7/metasploitable3`

[13] The prebuilt image used for testing can be downloaded from: `https://sourceforge.net/projects/metasploitable3-ub1404upgraded/files/`

[14] `https://www.virtualbox.org/`

By using relatively simple statistical models to represent text paragraphs, NLP allows for activities such as regression and classifications, hence generating additional smaller pieces of information that could be included in the knowledge base.

The approach followed for the textual analysis was generally similar among the different trials experimented with and consisted in:

1. Preparing the textual dataset: cleaning the target corpus by means of lemmatisation[15] and stemming[16] in order to simplify the set of words; purging the so-called stop-words, phrase particles that carry a low amount of information; and finally splitting the data into the training set and test set. The specific code for the textual preparation is the following:

```
import nltk
import gensim
from nltk.stem import WordNetLemmatizer, SnowballStemmer
nltk.download(’wordnet’)
nltk.download(’omw-1.4’)
stemmer = SnowballStemmer("english")

def lemmatize_stemming(text):
    return stemmer.stem(WordNetLemmatizer().lemmatize(text, pos=’v’))# Tokenize and lemmatize
def preprocess(text):
    result=[]
    for token in gensim.utils.simple_preprocess(text) :
        if token not in gensim.parsing.preprocessing.STOPWORDS:
            result.append(lemmatize_stemming(token))

    return ’ ’.join(result)
```

2. Choosing and training a model for feature extraction from textual paragraphs;

[15] `https://en.wikipedia.org/wiki/Lemmatisation`

[16] `https://en.wikipedia.org/wiki/Stemming`

3. Choosing and training a model for performing the chosen task of classification or regression;

4. Model's performance evaluation.

All the code necessary for this process was developed in the Python programming language[17], taking advantage of the existing open-source libraries dedicated to machine learning and data processing techniques such as SciKit-Learn[18], NLTK[19], Gensim[20], and Pandas[21]

The datasets required for the training and testing of the various models were either directly obtained from publicly available sources or constructed by refining or processing information obtained from public sources.

5.3.2 Classifying vulnerabilities from their description

The purpose of this task derives from the genericity and complexity of the textual descriptions produced by the tools evaluated in Section 5.2; being created in order to produce human-readable results, they present reports with a high level of verbosity, thus with a large amount of information that it is not needed. Moreover, while for a human reader may be a trivial task (although not in all cases), for an automated tool, it is not trivial to deduct the typology of the detected vulnerability by analyzing the textual description, hence machine learning strategies come into play with the objective of automatically extrapolating that relevant information. It is in fact considered that summarizing the attack information present in the description by means of high-significance labels could be significant for the purposes of risk assessment.

[17] `https://www.python.org/`

[18] `https://scikit-learn.org`

[19] `https://www.nltk.org/`

[20] `https://radimrehurek.com/gensim/`

[21] `https://pandas.pydata.org/`

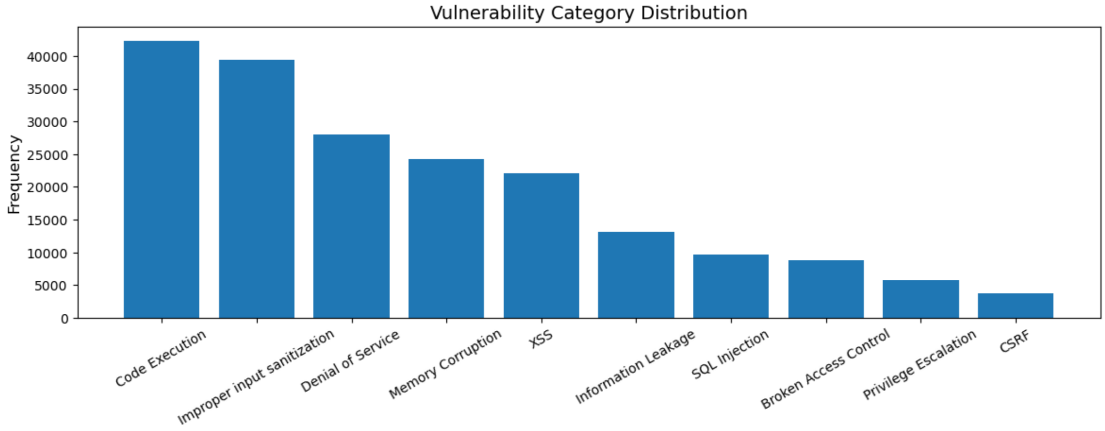


Figure 5.2: Number of vulnerability entries for each category.

Dataset

The data used to train and test the model has been constructed from `https://www.cvedetails.com`, a third-party website dedicated to hosting additional information and statistics about MITRE's CVE database entries[22], including but not limited to, particular labels describing the categories which each documented vulnerability belongs to. The collection of such data has been performed programmatically since it is not directly downloadable from the source; additionally, some similar categories were merged under the same in order to reduce the number of labels as well as increase the size of less represented types of vulnerability.

The final dataset contains 124741 vulnerability entries represented by their CVE-ID, a single textual description, and a list of categories to which it belongs. The descriptions, before pre-processing, show an average of 278.76 characters each. In total there are 10 different categories and each vulnerability may belong to more than one: *'Broken Access Control'*, *'CSRF'*, *'Code Execution'*, *'Denial of Service'*, *'Improper input sanitization'*, *'Information Leakage'*, *'Memory Corruption'*, *'Privilege Escalation'*, *'SQL Injection'*, *'XSS'* distributed as shown in Figure 5.2

[22] `https://cve.mitre.org/`

Models

For the vectorization phase of the process two different approaches have been tested: the first based on a Doc2Vec model[23] and the second based on a TF-IDF model[24]. While the first one should theoretically be a more advanced and elegant model based on word embeddings (i.e. a vector representation of a word in a multidimensional space, where the words that are closer in the space have similar meaning [12]), significantly better results in the classification were obtained modeling the descriptions via the frequency of the words they contained, multiplied by the inverse of the frequency of these terms in the whole corpus of descriptions. Hence, rather than having a vector representing a position of the term in an imaginary space, we employ a vector where each component represents one of the terms of the entire vocabulary, having as value the said product.

With regard to the model used for classification, the toolset provided by the SciKit-Learn library has been fully taken advantage of. Since the entries may be labeled with more than one category, the MultiOutputClassifier wrapper class[25] in order to train multiple classifiers able to recognize a single category of vulnerability, hence, by processing the data through each and every classifier it is possible to apply multiple labels to the same sample. The single class classifiers are instead based on the logistic regression models provided by the library[26].

Results

The validation of the models has been done with 10-fold cross-validation and determined some interesting results. The parameters of the cross-validation were set to use training sets of 70% of the size of the whole dataset leaving 30%

[23] `https://radimrehurek.com/gensim/models/doc2vec.html`

[24] `https://en.wikipedia.org/wiki/Tf%E2%80%93idf`

[25] `https://scikit-learn.org/stable/modules/generated/sklearn.multioutput.MultiOutputClassifier.html`

[26] `https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html`

of the entries as test data. The measures of performance for the multiclass model employed are:

$$Precision = \frac{TP}{TP + FP} \tag{5.1}$$

$$Recall = \frac{TP}{TP + FN} \tag{5.2}$$

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5.3}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.4}$$

With this method, an average F1 score of 0.97 was reached, which is consistent with the other measures obtained: a recall of 0.95 and a precision of 0.99. A separate consideration has to be made about accuracy; with the respective scoring function provided by the library, an entry is considered correctly classified if and only if all the labels predicted by the model exactly match the correct ones. However, with this behavior, partially correct classifications are ignored: for example, if an entry is labeled as *('Code Execution', 'Privilege Escalation')* but the models classify it as *('Code Execution')*, the prediction will be considered as erroneous. Nevertheless, this result is still able to provide significant information and should not be discarded. Therefore, by adjusting the metric in order to consider the errors on the individual labels instead of the whole data entry, it is possible to reach an accuracy of 0.99, with considerable results on individual classes, as shown in Figure 5.3. Given the particular relevance of these results, no other model has been tested other than the Logistic Regression classifier.

5.3.3 Deducing adversarial tactics from adversarial techniques and vulnerability descriptions

The main reason behind this task was trying to relate the discovered vulnerabilities with the possible tactics that they may enable and that an attacker may employ to exploit such weaknesses. The expected result is to obtain

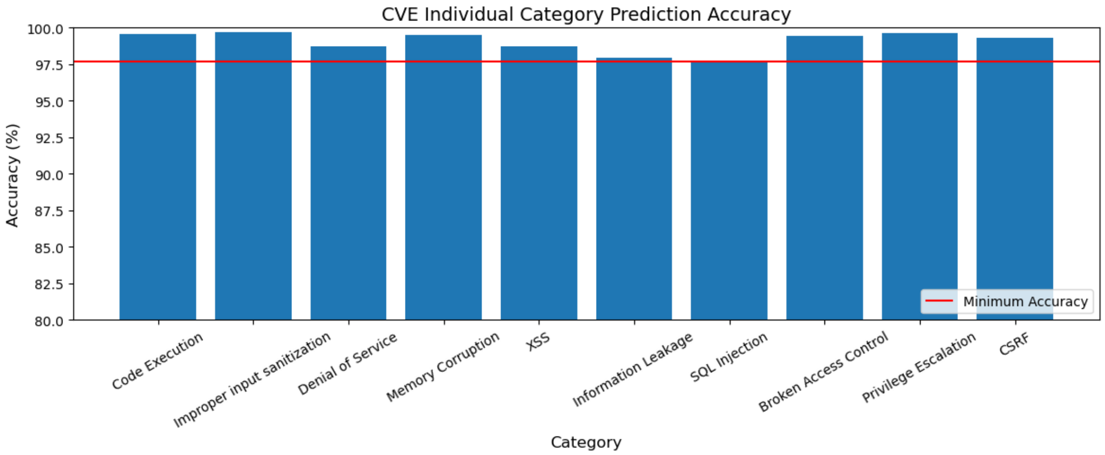


Figure 5.3: Prediction accuracy for each category of vulnerability.

additional information regarding the attack phase in which the vulnerability may become relevant, providing information about the priority of mitigation and possibly any insight on the potential impact of exploitation. Once again, the input data to be processed is the vulnerability description extracted by the vulnerability assessment reports; hence textual classification is performed. The adversarial tactics are selected according to the MITRE ATT&CK®[27] framework in order to produce results that are aligned with industry-level standards that are internationally recognized.

Dataset

The data used to train and test the models used for classification has been obtained from `https://attack.mitre.org/resources/working-with-attack/`, which contains useful downloadable resources of the framework. Among these resources, there are spreadsheets containing detailed textual descriptions of adversarial techniques that are associated with a category of adversarial tactics; since said techniques can be considered as a generalization of the description of a vulnerability, it is assumed that it could be similar to it both semantically and with respect to the used terminology, for this reason, is

[27] `https://attack.mitre.org/`

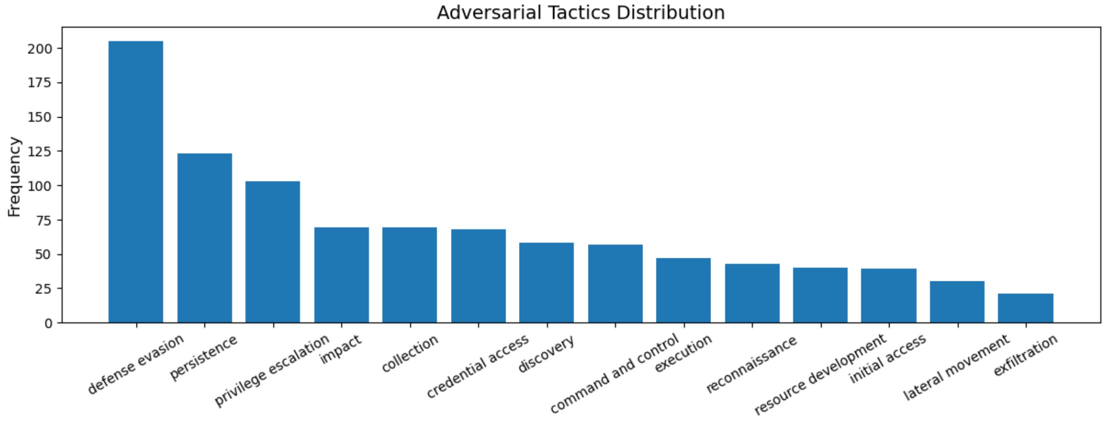


Figure 5.4: Number of adversarial techniques for each category of adversarial tactics.

supposed that a classification model trained on the techniques could also generalize on more specific vulnerability descriptions if the training set contains enough data.

After processing the original information, the final dataset comprises 778 entries representing different techniques, each representing an adversarial tactic defined by a name, a textual description, and a list of related adversarial tactics. The tactics identified by the framework are: *'collection', 'command and control', 'credential access', 'defense evasion', 'discovery', 'execution', 'exfiltration', 'impact', 'initial access', 'lateral movement', 'persistence', 'privilege escalation', 'reconnaissance', 'resource development'* distributed as shown in Figure 5.4

Models

Once again, the task has been developed by comparing different models both for vectorization and classification, employing the ones that led to better empirical results on the dataset. This time, with respect to textual vectorization, the Doc2Vec model is chosen since it appeared to bring significantly better results.

When dealing with classification instead, the simple statistical model used in the previous task and based on logistic regression was not producing significant results; hence, it has been decided to use the Multi-Layer Perceptron[28] model provided by the SciKit Learn library. The architecture of the primitive neural network was kept simple, using only a single hidden layer with 100 nodes with the hyperbolic tangent as a non-linear activation function. The reason behind those parameters can be found in the fact that they led to better empirical results and kept the training time low enough to validate the models efficiently.

Results

Again, the validation was performed with 10-fold cross-validation; although not significantly accurate, the model proved to be able to categorize the data entries decently. During this process, the dataset was split more using a 70-30% proportion of the training and test sets. Therefore, with the best-performing models, it was possible to obtain an average F1 score of 0.89, which is again consistent with the other measures: a recall of 0.88 and a precision of 0.91. Also, in this case, accuracy was lower, with a value of 0.81; when considering only exact multilabel predictions, setting the scoring function at the single label level, it increases to 0.98.

The results obtained for this type of classification are not particularly interesting; they indeed indicate that the model can discern different types of data entries, but the accuracy is too low to be considered a usable model. Moreover, when applying the model to text extracted from reports produced by the tools, which were not present in the dataset used for training and validation, the model appears to perform predictions that are not actually related to the paragraphs processed. This behavior may signify that the assumptions about similarities between the dataset and the real-world data may be wrong or that the dataset is not big enough for generalization.

[28] `https://en.wikipedia.org/wiki/Multilayer_perceptron`

5.3.4 Estimating real-world exploitability score from textual descriptions

The purpose of this experiment is to estimate the real-world likelihood of exploitation of the weaknesses identified by the vulnerability assessment tools to provide extremely important and measurable information about exposure to attacks of the target organization. In order to estimate such a complex indicator, the Exploit Prediction Scoring System (EPSS)[29] comes in help. This scoring system is generated by a model trained over several different features related to exploiting vulnerabilities in the wild, assigning to each CVE a value between 0 and 1 that indicates how likely to exploit the vulnerability is. Therefore the objective of this section will be to predict that same score only using the description of the vulnerability and eventually determine if the model is able to assign significant scores to more generic textual entries.

Dataset

The dataset used for training and evaluation is constructed starting from the one used in Subsection 5.3.2 with the addition of the EPSS data available at `https://www.first.org/epss/data_stats`. Merging the two data sources allows us to obtain 124741 vulnerability entries represented by their CVE-ID, a single textual description, and a score value in the range of 0-1.

As clearly evident in Figure 5.5, the vast majority of the CVEs present in the dataset have an assigned score between 0 and 0.1 (i.e. they have a very low likelihood of exploitation in the wild), this lack of uniformity needs to be accounted for when developing and training the models.

Models

Both the models presented in the previous sections for textual vectorization (i.e. TF-IDF and Doc2Vec) have been tested for this task, unfortunately,

[29] `https://www.first.org/epss/`

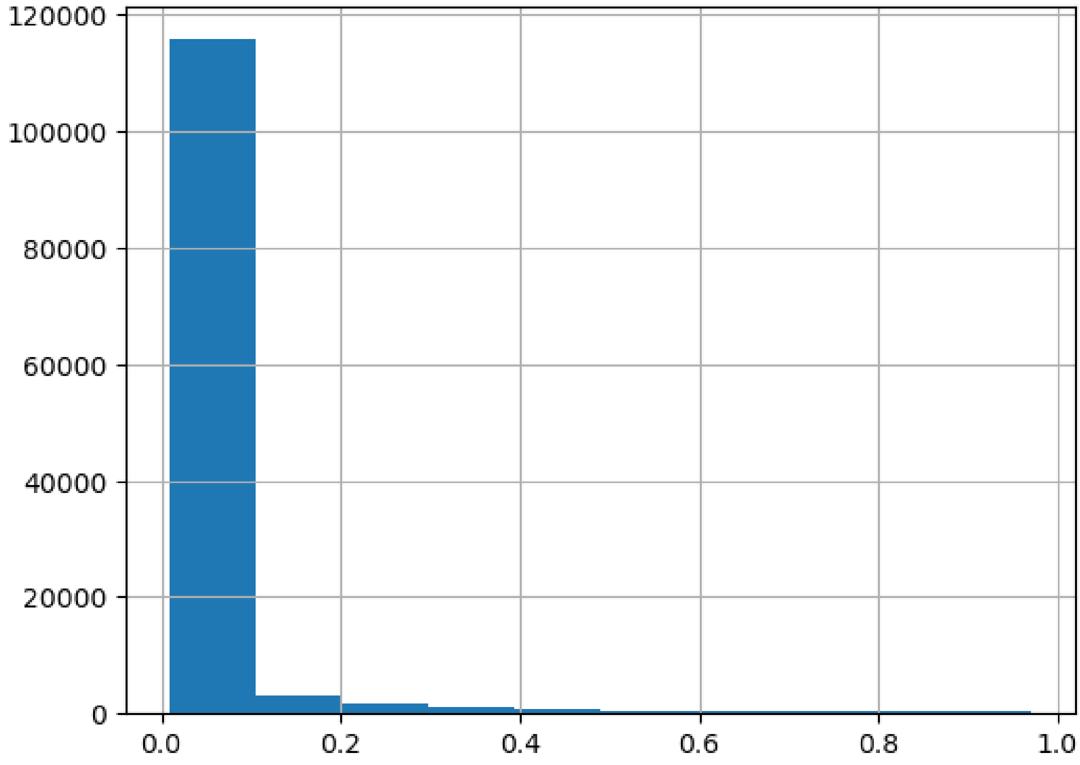


Figure 5.5: Distribution of the EPSS scores in the dataset.

neither of them showed significantly better performances than the other.

For what concerns the scoring estimation, two strategies have been tried:

- *Regression*: in order to obtain continuous values in the given range, different models for regression provided by the SciKit Learn library such as Linear Regressor[30] Multi-Layer Perceptron Regressor[31], and Random Forest Regressor[32] have been tested;

[30] `https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html`

[31] `https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html`

[32] `https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html`

- *Classification*: instead of predicting continuous values, the scores of the dataset were approximated and aggregated into the closest probability level, meaning that each vulnerability was assigned to a class of severity represented by an integer number. Ten, five, and three classes (i.e. levels from 0 to 9, from 0 to 4, and from 0 to 2, where a higher level indicates a higher likelihood of exploitability) classification problems were therefore solved instead of regression. The models tested for this purpose were: Linear SVC[33], Logistic Regression Classifier[34], and SGDClassifier[35] deployed in an One Vs Rest[36] approach in order to solve the multiclass problem; and a Random Forest Classifier[37].

Results

Unfortunately, neither of the proposed tasks achieved any significant results. Particularly worth mentioning is the lack of balancing of the scores among the dataset; keeping the entries unbalanced during training and validation produced values of accuracy (estimated via the coefficient of determination R^2 [38]) extremely promising, this result was, however, obviously incorrect because the model always predicted scores lower than 0.1 (even lower than 0.01) for each vulnerability description, this behavior, combined with the fact that entries with scores below 0.1 comprised 93% of the total entries, led to a particularly low average error of estimation.

Nevertheless, for our purposes, detecting vulnerabilities with high scores with reliability is vastly more important than assigning low scores, and none

[33] `https://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html`

[34] `https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html`

[35] `https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.SGDClassifier.html`

[36] `https://scikit-learn.org/stable/modules/generated/sklearn.multiclass.OneVsRestClassifier.html`

[37] `https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html`

[38] `https://en.wikipedia.org/wiki/Coefficient_of_determination`

of the models was able to do so.

Not even after uniformly balancing the dataset by randomly sampling from the original entries was able to improve the predictions, neither in the regression nor the classification task.

Cross-validation was not performed to evaluate the models since preliminary evaluations on the training set failed to achieve significant values of accuracy, with even lower values on the test set (i.e. none of the classification models was able to exceed a 70% accuracy while the regressors had R^2 values between negative and 0.40).

This lack of relevant results led to the conclusion that a textual description may not contain enough information to estimate a real-world exploitability score.

5.3.5 Classifying data breaches from their textual description

The last proposed task is another classification process. However, this time is not strictly related to the description of vulnerabilities but instead to the textual description of recorded attacks and data disclosures that afflicted worldwide organizations and companies. The purpose of this is to experiment with another possible source of information with a high level of verbosity, with the objective of extrapolating summarized key concepts. Despite not actually being validated, it is supposed that an improved version-derived model could be applied to reports of adversarial events that targeted the organizations.

Dataset

The dataset used for the purposes of this activity is a processed version of the data available from `https://privacyrights.org/data-breaches`, which is a list of different, publicly disclosed data breaches that happened from 2005 to 2019. After processing (i.e. removing entries without the fields relevant for the analysis), the dataset comprises 8220 breach entries that are

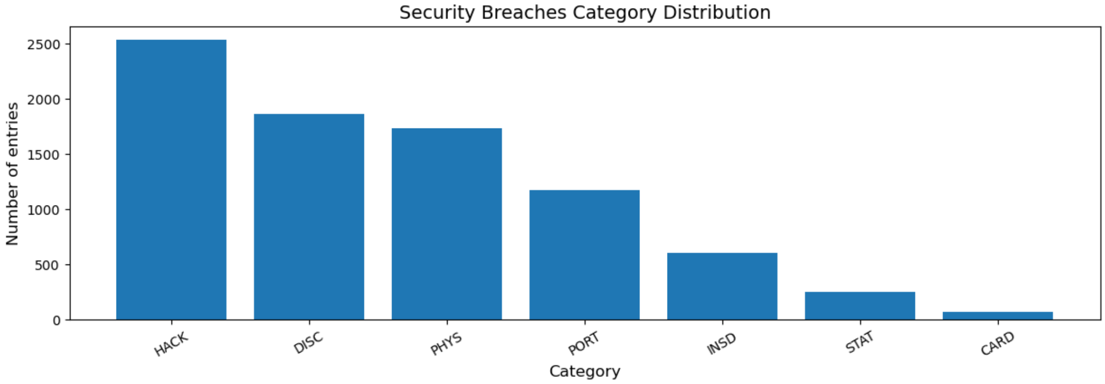


Figure 5.6: Distribution of entries of each category among the breaches' dataset.

represented by a textual description of the incident with an average length of 367 characters and a label describing the type of incident. The possible labels are:

- *CARD*: Fraud Involving Debit and Credit Cards Not Via Hacking (skimming devices at point-of-service terminals, etc.)
- *HACK*: Hacked by an Outside Party or Infected by Malware
- *INSD*: Insider (employee, contractor or customer)
- *PHYS*: Physical (paper documents that are lost, discarded, or stolen)
- *PORT*: Portable Device (lost, discarded, or stolen laptop, PDA, smartphone, memory stick, CDs, hard drive, data tape, etc.)
- *STAT*: Stationary Computer Loss (lost, inappropriately accessed, discarded, or stolen computer or server not designed for mobility)
- *DISC*: Unintended Disclosure Not Involving Hacking, Intentional Breach, or Physical Loss (sensitive information posted publicly, mishandled, or sent to the wrong party via publishing online, sending in an email, sending in a mailing, or sending via fax) [13]

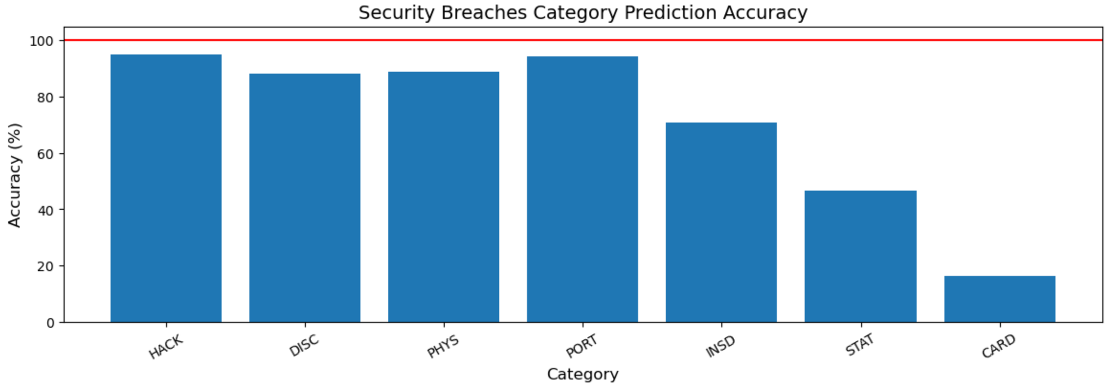


Figure 5.7: Prediction accuracy of the model for the different classes.

Models

The models adopted for this classification task are again selected through empirical experimentation and happen to be very similar to the ones employed in Section 5.3.2.

For textual vectorization, the model is based on TF-IDF, trained on the descriptions present in the dataset.

For classification, the model is once again a Logistic Regression model provided by the SciKit Learn library.

Results

Validation is performed once more with the technique of 10-fold cross-validation with a 70%-30% split for the training set and the test set. The evaluation is based on the accuracy measured in the predictions made by the model, which averages 0.88. The accuracy is also calculated per class, as shown in Figure 5.7.

Unsurprisingly, classes with a lower number of representing entries in the dataset appear to have lower average accuracy, worsening the general performance of the model. It is supposed that having a bigger, more uniform set of information may benefit the classification model.

Chapter 6

Reasoning Framework

This chapter partially acts as a wrapper around the knowledge gathered during the previous ones while simultaneously adding new technologies to the process of risk-relevant information collection to achieve better automation. After introducing the methodologies employed, the following sections describe the framework for information gathering that has been developed with its structure and functioning, as well as the data models it relies on.

6.1 Forward Reasoning and CLIPS

One important consideration is that cyber risk-related information is the product of a larger process of rules and conditions. Automating the production of information means creating a process that, given a small piece of data, can autonomously generate larger volumes of knowledge. Therefore sources of information alone are insufficient for this purpose: a framework orchestrating actions according to given rules is necessary to surround the knowledge base, allowing for its structured growth. The most straightforward and intuitive way of implementing such a framework in information systems is with automated reasoning methods, precisely forward reasoning.

This technique consists of the repeated application of modus ponens, a rule of inference based on implication (e.g., "P implies Q. P is true. Therefore

Q must also be true.") [14], to the available data. Instaurating dependencies between information and implication rules generates a model that can automatically produce new knowledge upon realizing given preconditions, thus achieving the required objective.

The implementation of such a model inside an information system is made possible by the software CLIPS, a tool and a programming language specifically built for the creation of expert systems (e.g., computer systems emulating the decision-making ability of a human expert[1]) that is open source, efficient and accessible. CLIPS main components are facts and rules; the firsts represent information, while the seconds indicate actions that will be executed when the facts satisfy the rule's preconditions, following an if-then paradigm.

6.1.1 CLIPSPy

Although conceptually powerful, the CLIPS environment and especially its CLI are limited when considering the need for interaction with external sources of information. Thankfully the open source project "CLIPSPy"[2] acts as a wrapper for the language, providing bindings to its functions inside a Python environment.

Another extremely important feature that this project adds to the language, is the possibility of invoking Python functions inside the CLIPS language as they were native functions; this feature allowed for numerous opportunities such as external API calls, HTTP requests, launching executable programs, and many more, all to increase the sources of information that can be integrated into the system.

[1] https://en.wikipedia.org/wiki/Expert_system

[2] https://github.com/noxdafox/clipspy

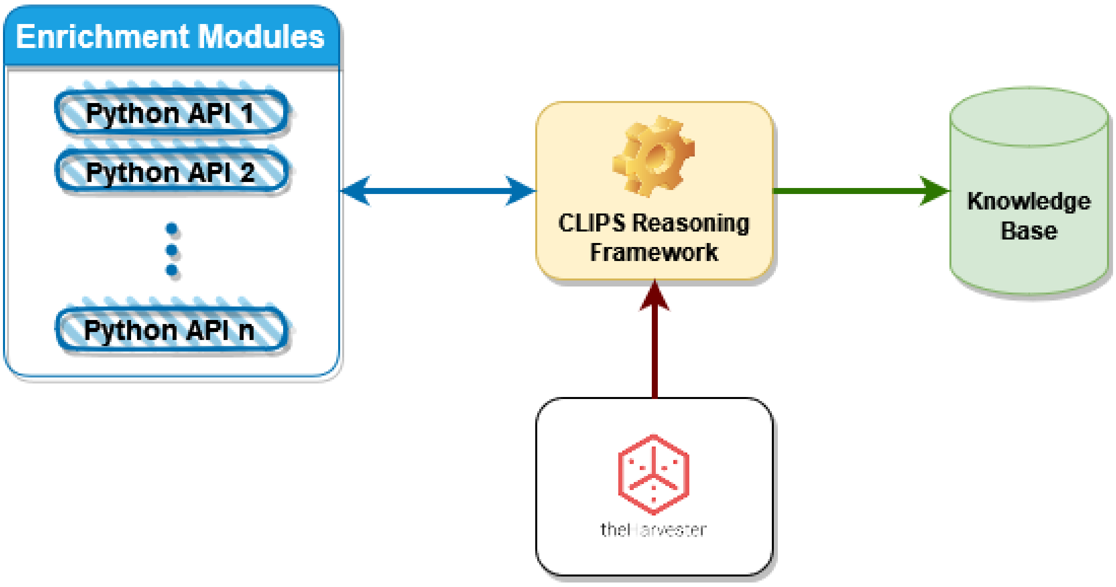


Figure 6.1: Schematization of the framework and its components.

6.2 Framework Structure

The information gathering framework built around the CLIPS reasoning system can be seen in a high-level overview in Figure 6.1. At the current state, the process relies on the open source tool theHarvester[3] which in turn uses the Shodan Search Engine[4] to gather the data necessary for the initial survey (i.e., related IP addresses and other subdomains) from the target initial domain. This approach has been chosen for simplicity of development and prototyping purposes, but it is assumed that the dependency on the said tool can be removed with minor efforts.

Once the initial information is obtained and trivially processed to be asserted into the CLIPS framework as facts or, simply put, information inside the knowledge base, the reasoning process can begin.

[3] `https://github.com/laramies/theHarvester`

[4] `https://www.shodan.io/`

6.2.1 Data Model

Figure 6.2 shows a representation of the simple initial data model used for representing the information that the current system processes and gathers. Structures for modeling data in the CLIPS language are called 'Templates.'

In the current implementation of the framework, the information collection process starts from the insertion of an *IP_Template* structure, which generally contains an IP address and a list of hostnames that resolve to said address (respectively, the address and hostnames fields of the template).

It is important to note that none of the fields of the templates are mandatory. While some are required to trigger the firing of other rules, theoretically, rules for the retrieval of specific fields may be written, assuming that different fields (or data entries) provide enough information. Therefore, facts that are incomplete at the beginning of the reasoning process may be modified and completed during the reasoning process.

Another important consideration has to be made about the relationships among the different data entities. The majority of links between pieces of information can be followed in both directions as explicitly represented in the *IP_Template* and *Domain_Template* relation. This means that the information can be obtained by forwarding reasoning. Still, also in a backward manner, starting from more specialized knowledge that could provide insight over the parent data entries, an inference approach called backward reasoning[5]. This aspect suggests that an adequate representation of the data model would be with technologies that rely on graphs; unfortunately, both backward reasoning and graph approaches were outside this thesis's scope.

6.2.2 CLIPS Rules

Another significant component of the reasoning framework is the set of CLIPS inference rules that guide the collection process. Written in the

[5] https://en.wikipedia.org/wiki/Backward_chaining

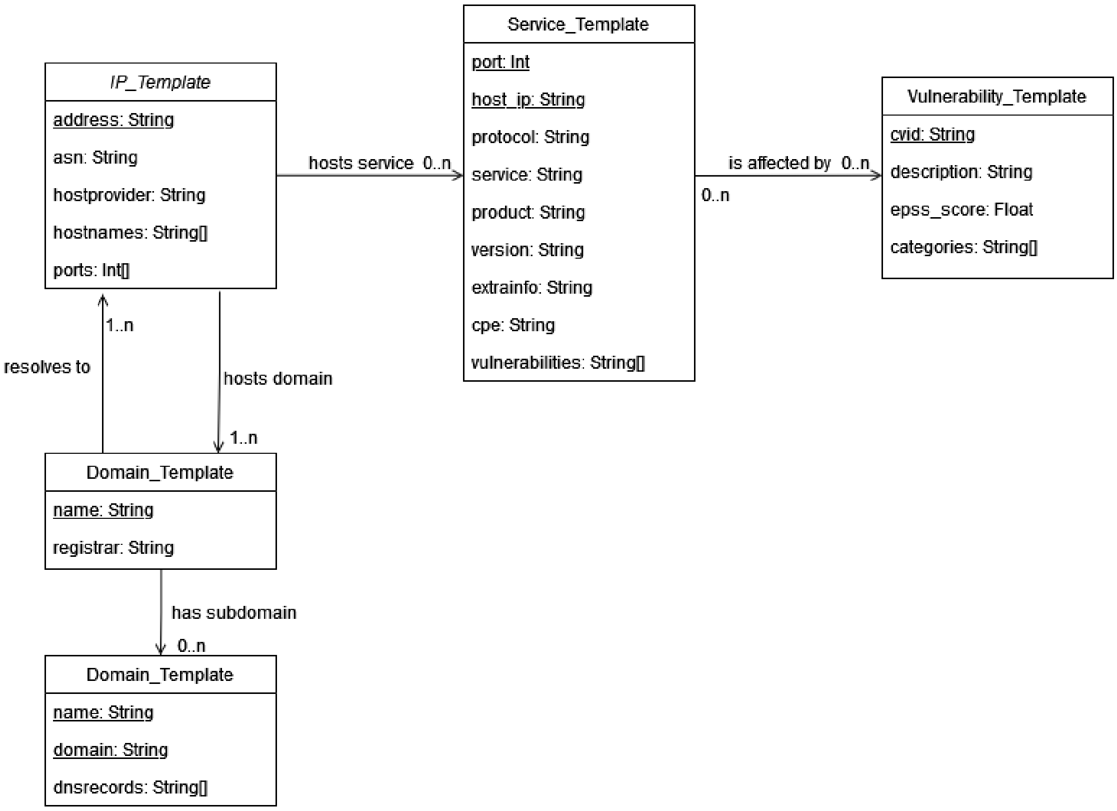


Figure 6.2: Diagram of the data entries initially modeled in the framework.

CLIPS language, these rules activate depending on the set of facts present in the system and execute different actions, including inserting new facts into the knowledge base.

```
(defrule SCANRULE
  (declare (salience 51))
  ?ipt <- (IPTEMPLATE (address ?address) (ports $?ports))
  (test (neq ?address ""))
  (test (eq (length$ ?ports) 0))
  =>
  (printout t "Portscan for ip: " ?address crlf)
  (bind $?newports (get_ports ?address))
  (modify ?ipt (ports $?newports))
)
```

The above code is an example of a CLIPS rule for executing a port scanning action against the IP addresses in the system. The rules generally follow the same structure identified by two parts separated by the arrow symbol “`=>`”: the Left Hand Side or, in short, LHS being the lines from 2 to 4 and the Right Hand Side or RHS, which corresponds to lines from 6 to 8 in the example.

The LHS defines the patterns (i.e., preconditions) that a fact needs to satisfy to trigger the rule; in the example, the required fact needs to be an *IP_Template* with a non-empty address that has the multislot (i.e., a field that can contain multiple values, similar to an array) field *ports* empty, that conceptually indicates an IP address that has not yet been scanned for port discovery. Line number 2 defines the salience of the rule, being the priority of execution of said rule: the rule with higher salience is always executed first.

The RHS instead defines the actions to be performed when the rule is executed; in the case of the example, the rule first outputs a message to the user announcing the execution, then calls the enrichment function *get_ports* passing as an argument the field *address* and finally updating the *IP_Template* fact setting the discovered ports in the relative field.

Lastly, it is essential to note that the same, unmodified fact will not trigger the same rule more than once; this avoids the execution of infinite loops during the execution of the rules agenda.

At the current state, a total of 10 initial rules have been developed and tested.

6.2.3 Enrichment Modules

The last fundamental components of the system are those responsible for collecting information from external sources. These components have been named ‘Enrichment Modules’ since they aim to develop the system’s knowledge base by providing new information based on the previously obtained, thus enriching the actual knowledge base.

As briefly mentioned in Section 6.2.2, these modules are implemented as internal functions callable from the CLIPS environment and, in particular, in the RHS of the rules. However, this behavior is made possible by the CLIPSPy bindings: the actual implementation of the modules is made using functions written in the Python language, whose references are then imported into the CLIPS environment allowing for their usage.

This approach proved ideal and significantly simplified the development of rules and enrichment modules by introducing a well-established, fully featured, and supported language like Python 3 into the development stack. As a consequence of this, many features were made possible to be implemented in the system, from simple HTTP requests to web API endpoints, to the execution of local programs such as NMAP through its Python library, as well as the integration of the machine learning models developed in Section 5.3 for vulnerability classification.

Chapter 7

Conclusions

The problem of automated information gathering for cyber risk assessment purposes proved to be a very broad and complex theme to be studied and analyzed. The diverse nature of the process components, as well as the conceptual cut of the problem itself — with its general lack of concrete and independent methodologies, definitions, data, and instruments — resulted in a particularly time-consuming research process that limited the depth of investigation in the various components. This is further emphasized by the restricted availability of literature dealing with this subject. Nevertheless, exciting results have still been obtained despite the difficulties encountered.

The first successful step to better understand the problem was investigating the standard procedures in the cyber risk management processes. This step was done via the study of the industry-level standards. This process led to the identification of high-level concepts that, with more work, can be quantified and applied to organizations and companies to provide measures for their cyber risk exposure, possibly in an automated manner.

Another result derived from the best practices study was the identification of existing approaches related to cyber risk assessment that were well suited towards automation or that were already partially automated. Studying the various tools used for vulnerability assessment purposes gave an insight into what information was required to estimate the risk derived from vulnerable information systems; this research also revealed various techniques used for said purposes that could be automated and integrated into other information-gathering systems. Additionally, this investigation led to the experimentation with different machine learning models to improve and enrich the data that specific tools produced. This approach resulted in a model that accurately classified vulnerabilities based on their textual description.

The last result produced during the thesis, but probably the most relevant one, is the development of an expert system prototype built around the knowledge obtained throughout the research process. Based on forwarding reasoning, the system at the current stage can automatically retrieve basic information about specified targets but showed considerable possibilities for improvements and expansion, proving the existence of numerous options for an automatic process of cyber risk-relevant information gathering.

7.1 Future work

Several have been identified during this thesis' development regarding future work opportunities. Beginning with the theoretical concepts identified in Chapter 4, it is assumed that the further elaboration of such ideas could produce interesting measures and indexes that could properly quantify certain risk factors and consequently be combined to assess a cyber risk value. The evaluation of said measures is expected to be done both via interactive and human-dependent manners (i.e., company surveys), as well as with the employment of automatic tooling when the necessary information could be derived without the need for human intervention, for this last approach; however, further investigation is required.

Apart from this, the machine learning approaches developed in Chapter 5 could be significantly improved by dedicating more effort to the data quality, training processes, and research about the underlying models used. Other enrichment strategies apart from categorization and regression could also be practiced, and instead of Natural Language Processing over textual data, different approaches could be also tested.

But the most significant improvements could be carried out toward the reasoning framework developed in Chapter 6. The current version of the expert system consists of a prototype with a limited set of information types and rules of inference, as well as a restricted group of enrichment functions. With new data templates, the number of regulations will also increase and, consequently, the number of required enriching modules. Possible integration with the surveys and measures derived from the concepts of Chapter 4 are also theorized, with a particular interest in company questionnaires that have questions with dependency relations. Lastly, the rules could also be expanded using backward reasoning strategies instead of sole forward reasoning, although more research is required to reach a proper design.

Bibliography

[1] Joint Task Force Transformation Initiative Interagency Working Group. *Managing Information Security Risk. Organization, Mission, and Information System View*. Tech. rep. NIST Special Publication (SP) 800-39. Gaithersburg, MD: National Institute of Standards and Technology, 2011. DOI: 10.6028/NIST.SP.800-39.

[2] Joint Task Force Transformation Initiative Interagency Working Group. *Guide for Conducting Risk Assessments*. Tech. rep. NIST Special Publication (SP) 800-30, Rev. 1. Gaithersburg, MD: National Institute of Standards and Technology, 2012. DOI: 10.6028/NIST.SP.800-30r1.

[3] Joint Task Force Transformation Initiative Interagency Working Group. *Security and Privacy Controls for Information Systems and Organizations*. Tech. rep. NIST Special Publication (SP) 800-53, Rev. 5, Includes updates as of Dec. 10, 2020. Gaithersburg, MD: National Institute of Standards and Technology, 2020. DOI: 10.6028/NIST.SP.800-53r5.

[4] Financial Services European Commission Directorate-General for Financial Stability and Capital Markets Union. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014 and (EU) No 909/2014*. 2020. URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex:52020PC0595.

[5] *Digital finance: Provisional agreement reached on DORA*. [Online; accessed 18. Aug. 2022]. Aug. 2022. URL: https://www.consilium.europa.eu/en/press/press-releases/2022/05/11/digital-finance-provisional-agreement-reached-on-dora.

[6] *Systemic cyber risk*. Feb. 2020. URL: https://www.esrb.europa.eu/pub/pdf/reports/esrb.report200219_systemiccyberrisk~101a09685e.en.pdf.

[7] European Central Bank. *TIBER-EU*. 2018. URL: https://www.ecb.europa.eu/pub/pdf/other/ecb.tiber_eu_framework.en.pdf.

[8] Frank Cremer, Barry Sheehan, Michael Fortmann, Arash N. Kia, Martin Mullins, Finbarr Murphy, and Stefan Materne. «Cyber risk and cybersecurity: a systematic review of data availability». In: *Geneva Pap. Risk Insur. Issues Pract.* 47.3 (July 2022), pp. 698–736. ISSN: 1468-0440. DOI: 10.1057/s41288-022-00266-6.

[9] Mark Mateski, Cassandra M. Trevino, Cynthia K. Veitch, John Michalski, J. Mark Harris, Scott Maruoka, and Jason Frye. *Cyber Threat Metrics*. 2012. URL: https://irp.fas.org/eprint/metrics.pdf.

[10] *2022 Ponemon Cost of Insider Threats Global Report | Proofpoint US*. [Online; accessed 28. Sep. 2022]. May 2022. URL: https://www.proofpoint.com/us/resources/threat-reports/cost-of-insider-threats.

[11] *Vulnerability Scanning Tools*. [Online; accessed 19. Aug. 2022]. Aug. 2022. URL: https://owasp.org/www-community/Vulnerability_Scanning_Tools#.

[12] Contributors to Wikimedia projects. *Word embedding - Wikipedia*. [Online; accessed 16. Sep. 2022]. Aug. 2022. URL: https://en.wikipedia.org/w/index.php?title=Word_embedding&oldid=1106009992.

[13] *Data Breaches | PrivacyRights.org*. [Online; accessed 22. Sep. 2022]. Sept. 2022. URL: https://privacyrights.org/data-breaches.

[14] Contributors to Wikimedia projects. *Modus ponens - Wikipedia*. [Online; accessed 20. Aug. 2022]. June 2022. URL: https://en.wikipedia.org/w/index.php?title=Modus_ponens&oldid=1094077354.