Computer Science and Information Systems 23(1):343–368 https://doi.org/10.2298/CSIS251030011G

DICYME: Dynamic Industrial Cyber Risk Modelling Based on Evidence*

Javier García-Ochoa[1], Jaime Rueda[1], Rubén R. Fernández[1], Alberto Fernández-Isabel[1], Isaac Martín de Diego[1], Emilio L. Cano[1], Romy R. Ravines[2], Ovidio López Espinosa[2] and Jaume Puigbó Sanvisens[2]

[1] Rey Juan Carlos University
Department of Computing Science & Statistics, ETSII
C/ Tulipán, s/n, 28933, Móstoles, Madrid (Spain)
{javier.garciaochoa, jaime.rueda, ruben.rodriguez, alberto.fernandez.isabel, isaac.martin, emilio.lopez}@urjc.es

[2] DeNexus Inc.
Boston, United States
{rr, ol, jp}@denexus.io

Abstract. The accelerated pace of digital transformation has significantly reshaped the cybersecurity domain, fostering an interconnected ecosystem in which cyber threats have expanded in both their complexity and scope. Traditional cybersecurity methods are increasingly inadequate for addressing the rapidly evolving threat landscape, emphasizing the critical need for intelligent, adaptive, and proactive defensive strategies. This study introduces Dynamic Industrial Cyber Risk Modelling Based on Evidence (DICYME), a comprehensive system that integrates diverse analytical techniques to identify patterns and characteristics that reveal emerging threat trends, enabling organizations to proactively defend against potential future attacks. Beyond threat detection, DICYME operates as a pipeline that retrieves data from diverse cyber incident reports, specialized databases, and other relevant sources of cyber-related information, applies specialized techniques for victim identification, indicator computation, threat actor profiling, Common Vulnerability and Exposure (CVE) relationship mapping, and ultimately performs the Cyber Risk Quantification (CRQ). This final stage represents the system's most distinctive contribution, as it translates complex analytical outputs into actionable risk insights, empowering organizations to make informed strategic decisions in the face of evolving cyber threats. Alternatively, the system implements an automatic workflow that constructs new datasets of compromised entities, enabling these datasets to be used by all components of the system. Experiments on real cyber incident datasets demonstrate the system's ability to automatically construct high-quality victim profiles and estimate annualized financial risk, offering a scalable and data-driven approach for proactive cybersecurity management.

Keywords: Cyber risk quantification, Machine Learning, Large Language Models, Indicators, Firmographics, Threat actors, Vulnerabilities.

* This is an extended and updated version of the paper "Dynamic Industrial Cyber Risk Modelling based on Evidence (DICYME)" presented at the 10th Jornadas Nacionales de Investigación en Ciberseguridad (JNIC 2025), Zaragoza, Spain.

1. Introduction

The exponential growth of digital transformation has fundamentally reshaped the global cybersecurity landscape, creating an interconnected ecosystem in which cyber threats have evolved in both sophistication and scale. The accelerated migration of individuals and organizations toward digital environments, particularly intensified by digital transformation, has expanded the attack surface and created new vulnerabilities across multiple technological layers, including cloud computing, Internet of Things (IoT) devices, social media platforms, and cryptocurrency systems [17]. This dynamic environment has rendered traditional, static cybersecurity strategies increasingly inadequate, as they often fail to anticipate emerging threats or provide real-time insights into the evolving risk profile of an organization.

Organizations face the dual challenge of identifying potential threats and quantifying their impact in a timely manner. Conventional cyber risk assessments rely heavily on manual processes and historical data, which are often incomplete, delayed, or incompatible across sources. This results in a slow, resource-intensive evaluation that struggles to capture the rapidly changing threat landscape and the adaptive and dynamic structure of organizations. Moreover, many existing approaches treat vulnerability and threat intelligence in isolation, failing to integrate multiple data streams into a unified, actionable risk framework.

Given the constant threats that can compromise enterprises across all sectors, there is a critical need for intelligence-driven cybersecurity measures that go beyond traditional methods. Integrating Machine Learning (ML) and Artificial Intelligence (AI) methodologies [10] allows organizations to adopt proactive, adaptive strategies capable of real-time risk quantification and data-driven decision-making. By leveraging the temporal patterns of cyber incidents through techniques such as time-series analysis, it becomes possible to anticipate future attack occurrences and detect anomalies that may indicate imminent threats [13]. This combination of automation, advanced analytics, and predictive modeling provides a foundation for more efficient, accurate, and dynamic cyber risk management.

To address these limitations, this study presents DICYME, a comprehensive system that combines automatic intelligence extraction, advanced ML models, and real-time risk quantification. DICYME leverages both public and private datasets, dynamically incorporating new vulnerabilities, emerging attack techniques, and threat actor behaviors. The system translates this intelligence into quantitative indicators and probabilistic scores, which are then used to estimate the likelihood, frequency, and potential impact of cyber incidents, including both primary and secondary losses.

A key feature of DICYME is the usage of Large Language Model (LLM) to generate interpretable explanations and actionable recommendations, helping stakeholders understand the rationale behind risk assessments and mitigation strategies. Additionally, it is adaptable to different stakeholder needs, supporting interactive visualization and reporting tools suitable for technical teams, executive decision-makers, insurers, and reinsurers, providing tailored insights that facilitate informed, data-driven decisions.

By integrating real-time data extraction, predictive modeling, and dynamic risk quantification, DICYME overcomes the fragmented and static nature of traditional cyber risk assessments. It provides a unified, continuously updated view of cyber exposure, enabling organizations to proactively detect emerging threats, prioritize security investments, and

respond effectively to the evolving digital threat landscape. This integrated approach represents a significant advancement in intelligence-driven cybersecurity, bridging the gap between automated data collection, analytical modeling, and operational decision-making.

The rest of this paper is organised as follows. Section 2 provides the background and related work, covering risk analysis, actor profiling, and CVE relationships. Section 3 introduces the proposed system, including automated data extraction, risk modelling, and quantification, with support from LLMs and AI tools. Section 4 presents a series of experiments evaluating the data gathering process of the system and the CRQ model. Section 5 discusses the lessons learned, highlighting the strengths, limitations, and practical insights gained from the deployment of the system. Finally, Section 6 concludes the paper and outlines potential directions for future research.

2. Background

The cybersecurity landscape has evolved into a complex ecosystem in which cyber incidents, risk analysis, threat actors, and vulnerabilities are interconnected elements that collectively shape organizational security postures [28]. Understanding these relationships is crucial for developing comprehensive defensive strategies and predictive models [27], as organizations face increasingly sophisticated and frequent attacks that exploit the interdependencies among these elements. Recent research emphasizes the need for holistic approaches that integrate cybersecurity risk management with strategic management practices, leveraging AI and ML techniques with traditional cybersecurity frameworks [19, 27].

In recent years, cyber incidents have been steadily increasing, becoming a common problem affecting organizations across multiple domains. Simultaneously, these incidents have grown in complexity [26], compromising a wide range of companies from different sectors, including sensitive areas such as healthcare and finance. According to Hackmageddon, the number of recorded cyber incidents continues to rise yearly, reaching $4,128$ events in 2023, representing a 35% increase compared to 2022 and a markedly higher growth relative to earlier years [22].

This trend has motivated companies and institutions to develop various techniques and approaches to collect, analyze, and explain cyber incident data, as well as to model the associated risks. At present, there are numerous repositories of cyber incidents that publish detailed information about reported cases, such as the European Repository of Cyber Incidents (EuRepoC) [5] and the Cyber Events Database of the Center for International & Security Studies at Maryland (CISSM) [1]. The reporting of such incidents, which may affect any type of organization, enables the identification of potential attack trends and patterns. This, in turn, supports the anticipation of future threats and provides valuable insights for security teams, decision makers, and other relevant parties who access and analyse these repositories.

2.1. Risk analysis

Based on the foundation of incident data collected in repositories, organizations require systematic approaches to assess and quantify cyber risks. Companies and institutions often conduct cyber risk analyses [24] to identify threat actors that are likely to carry out

attacks, as well as to quantify the probability and frequency of such attacks. Based on the results of this analysis, organizations can make informed decisions about whether to invest in security, what types of security measures to prioritize, and which techniques to adopt to protect themselves against common threats that may affect their specific business sector.

By leveraging databases that host records of cyber incidents affecting companies and institutions, it is possible to conduct comprehensive risk analyses. Probabilistic and statistical analyses [24, 30] are the most commonly applied approaches for assessing these risks, and they can be combined to develop descriptive models. Furthermore, it is possible to build models that function as risk management tools, enabling organizations to reduce the likelihood of threats and their potential impact.

However, these approaches are often based on historical data and present limitations in anticipating novel or rapidly evolving threats. Another important challenge in cyber risk analysis is the limitations or absence of certain relevant data. In many cases, the available repositories or databases do not provide different coverage for specific types of risks, making it necessary to rely on external or private data sources to conduct a more comprehensive analysis. Furthermore, it is mainly a manual process, performed from time to time, requiring effort from analysts that increases cost and reduces responsiveness to emerging threats [25]. In this context, our proposal introduces a system that integrates automated data extraction with real-time risk quantification, providing continuously updated and dynamic assessments tailored to the current threat landscape.

2.2. Threat actors analysis

While risk analysis provides a quantitative framework for understanding cyber threats, identifying and characterizing the specific actors behind these threats represents another critical dimension of cybersecurity research. Threat actors are individuals or organized groups responsible for causing cyber incidents in companies [6] leading to financial losses, reputational damage, sensitive data breaches, and broader security implications. They are often identified to conduct analyses that allow organizations to understand attack patterns, predict future threats, and develop targeted defense strategies.

The identification of threat actors often lacks consensus, as disagreements may arise regarding who should be considered an attacker. In many cases, the analysis tends to prioritize quantitative aspects [31], such as counting how often different threat actors are mentioned in reports, over qualitative considerations such as comparative or contextual evaluation. This quantitative focus, while providing measurable insights, may overlook important behavioral patterns and motivational factors that threat actor activities.

Several studies and frameworks have been proposed to address this issue by collecting data from reliable sources, enabling more comprehensive analyses to detect attack patterns, identify countries that are more susceptible to specific actors, and determine which types of organizations are most frequently targeted. These frameworks contribute to a more detailed understanding of the threat landscape and support the development of targeted mitigation strategies for specific actors.

Our approach offers a systematic way to categorize threat actors using publicly available data, while remaining adaptable to incorporate proprietary intelligence feeds from providers such as CrowdStrike or FireEye. Importantly, the relevance and impact of a

specific threat actor are contextual and victim-dependent: an actor predominantly targeting organizations in the United States or Canada may pose minimal risk to a company in Madagascar, but significant risk to a U.S.-based critical infrastructure operator. By integrating geographic, sectoral, and actor-specific prevalence data, our method provides dynamic, context-aware threat assessments that reflect both the actor's capabilities and the characteristics of the potential victim.

2.3. CVE relationship

Complementing the understanding of threat actors and their behaviours, vulnerability analysis provides a technical foundation to understand how attacks are executed and systems are compromised. CVEs represent a standardised framework for cataloguing publicly known security flaws that can affect organisations across all sectors [32]. When successfully exploited, these vulnerabilities enable threat actors to compromise systems, resulting in unauthorised data access, operational disruptions, and significant financial impact.

The association of vulnerabilities with affected systems is based on structured mapping processes that link CVEs to standardised repositories. Each vulnerability is assigned to the National Vulnerability Database (NVD), which supplements basic CVE information with impact assessments, references to vulnerable products via Common Platform Enumeration (CPE), and classification of the underlying weaknesses through Common Weakness Enumeration (CWE) [35]. This structured approach enables consistent vulnerability tracking and supports the integration of threat intelligence into organisational risk frameworks.

Recent advances in vulnerability relationship modelling have emphasised graph-based representations to capture complex interdependencies. Knowledge graph approaches [32, 37] enable the discovery of non-obvious connections between vulnerabilities, affected products, and exploitation patterns, supporting more accurate threat predictions. ML techniques, particularly graph neural networks [2, 38], have demonstrated substantial improvements in identifying vulnerable code patterns and predicting missing relationships within vulnerability databases. New research has extended these methods by incorporating Natural Language Processing and reasoning capabilities from LLMs to extract and integrate information from unstructured threat intelligence sources [16].

Despite these advances, vulnerability databases continue to face fundamental limitations. Significant processing delays result in incomplete metadata for newly disclosed vulnerabilities, while inconsistent weakness classifications reduce the reliability of automated analysis tools. Studies have documented substantial backlogs in vulnerability analysis and highlighted discrepancies in how CVEs are assessed and used in security research. Addressing these challenges requires continued development of automated enrichment techniques, cross-source validation mechanisms, and predictive models capable of operating effectively with incomplete or evolving vulnerability information.

3. Proposal

In this section, we present the proposed system, called DICYME, which focuses on automating cyber risk quantification and management in operational technology (OT) environments. The main objective of this system is to design and implement a system capable

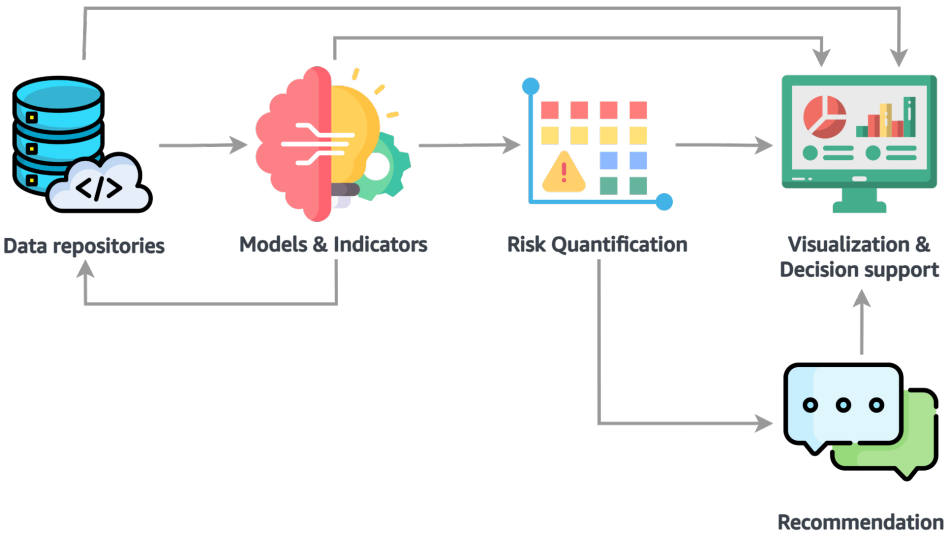

Fig. 1. High-level architecture of the DICYME complete system

of addressing the current challenges in modelling cyber risks within industrial infrastructure. The proposed solution relies on evidence-based analysis, integrating data from multiple sources to dynamically estimate the probability and impact of potential cyber incidents. To achieve it, the system incorporates a set of cyber risk indicators, which together with the data contribute to the simulation and quantification model. In addition, LLM to support explainability and enhance user understanding while providing recommendations in natural language. Finally, all components are brought together within a visualization platform that delivers a continuous, clear, and actionable view of cyber risk levels, designed to support informed decision-making in critical environments.

Figure 1 illustrates the architecture of the system, which is structured into five main parts. Automated data extraction gathers and preprocesses heterogeneous information from both external and internal sources, such as threat intelligence feeds and internal security telemetry. This raw evidence feeds the indicators block, where relevant cyber risk metrics are computed and normalised to reflect dynamic aspects such as visibility, reputation, or the threat actors landscape. These indicators are then aggregated in the cyber risk quantification module, which applies models to estimate the probability and impact of incidents. The outcomes, together with the underlying models and source data, are presented in the visualization and decision support layer, providing interpretable dashboards and analytical tools that allow stakeholders to explore and understand risk. Finally, the recommender component not only guides users in interpreting the visualized data and models, but also actively supports decision-making by proposing adjustments to the quantification process, suggesting alternative data inputs, and executing the risk models with multiple scenarios to provide actionable and feasible recommendations.

In addition, Figure 2 provides a more detailed view of the system, explicitly mapping all the elements that contribute to CRQ and the relationships among them. The figure distinguishes between the different data repositories, the models and indicators built on

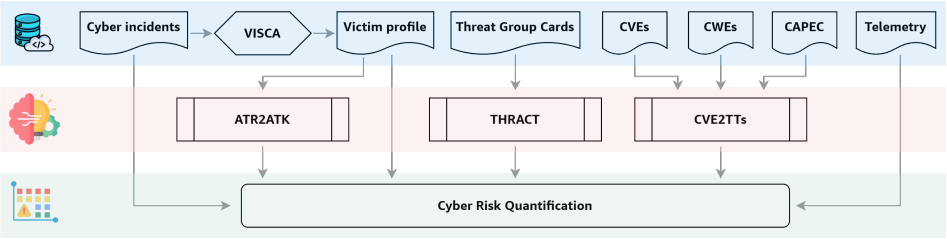

Fig. 2. Structural relationships within the Cyber Risk Quantification system

top of them, and the final CRQ process. This separation highlights how raw evidence is progressively transformed into structured knowledge and then into quantifiable indicators of risk, while also allowing certain datasets to be used directly in the quantification process when relevant, for instance to identify trends, contextual factors.

3.1. Automated data extraction

The automated data extraction module was designed to systematically collect and consolidate heterogeneous cyber risk information from multiple open sources. Specifically, it integrates reports of cyber incidents from six publicly available databases to provide a structured view of the threat landscape. To complement incident-level data, the module also incorporates victim profiling information, both with manual identification which is then empowered through Victim Insight System for Cyber Attacks (VISCA), a multi-agent LLM-based architecture that identified affected organizations and retrieves and processes its firmographic attributes.

In addition to incident and victim data, the module gathers technical Cyber Threat Intelligence (CTI) such as up-to-date CVEs with their associated vulnerability types, Tactics, Techniques and Procedures (TTPs) from the MITRE ATT&CK framework, as well as structured attack patterns from Common Attack Pattern Enumeration and Classification (CAPEC). Furthermore, it integrates information on threat actors from publicly available Electronic Transactions Development Agency (ETDA) Threat Group Cards [4], linking adversarial groups to their targeted victims and contextual attributes. It also incorporates internal telemetry data, capturing Operational Technology (OT)-specific information such as asset counts, exposed vulnerabilities, and network-level observations across Purdue layers. By consolidating organizational, incident-specific, adversarial, and technical dimensions, the module creates a unified knowledge base that supports the advanced modelling of cyber risks in subsequent stages.

Cyber incidents Cyber incident data were initially extracted from publicly accessible and free sources, including the TI Safe Incident Hub [34], Industrial Control System Security, Threats, Regulations, Incidents, Vulnerabilities provided by Experts (ICS STRIVE) [9], KonBriefing [12], Cyber Events Database of the CISSM, Hackmageddon [23] and EuRepoC. The extracted records were stored along with the extraction date and the source of the origin. Subsequently, the data underwent transformation and cleaning processes to achieve a more standardized format suitable for analysis. The transformation primarily

consists of restructuring the raw data into a tabular form, where each row represents a cyber incident and each column corresponds to an attribute of that incident. Data cleaning is performed to improve data quality by removing invalid entries or correcting formats, such as normalizing date representations. Additionally, cleaning involves harmonizing categorical values, for example, standardizing the country of the victim or the type of attack.

Finally, the sources were converted into a uniform tabular format to enable cross-source comparison of incidents, support exploratory data analysis across all repositories, facilitate incident deduplication, and allow for higher-level transformations.

Victim profile The Victim profile dataset was collected in the DICYME system to support the development of the Attractiveness concept (see Section 3.2), which delves with the possession of properties or behaviours in entities that may capture the focus of potential adversaries [7, 8]. The dataset includes uniquely identifiable entities that have been reported as victims of confirmed cyber incidents from the Cyber incidents dataset (see Section 3.1), manually selected and validated by a threat intelligence analyst. In addition, similar entities by country and industry category without known incidents were included to provide more context for a comparative analysis.

For each entity, the collected information can be organised into three groups that correspond to the main components of the Attractiveness model. The first group, basal attractiveness, refers to inherent static firmographic features, which includes the country, sector category, revenue, earnings, publicly traded status, number of employees, and profitability. The second group, online reputation, captures dynamic factors linked to human-generated content, social media interactions, and shared experiences related to the entity. It is computed through engagement (E), defined as the ratio of interaction (I) to reach (R) [3], where reach indicates the number of individuals who view or mention a publication and interaction represents those who actively respond to it. Engagement is assessed by distinguishing whether the content originates from the entity itself (assumed predominantly positive) or from external users, where sentiment analysis is performed to capture positive, neutral, or negative perceptions [7]. The third group, potential victimisation, represents another dynamic factor and includes the number of appearances of the entity name in dark web leaks and the number of visible devices connected to the Internet. These three groups are combined to produce a final structured dataset, containing a large amount of information, including features that are complex to extract, all referenced to each entity.

Victim Insight System for Cyber Attacks (VISCA) The VISCA model is an agent-based framework developed within the DICYME system to improve and extend the Victim profile dataset by automating and scaling the identification of victim entities and the fusion and normalization of heterogeneous data. This model takes the description of a cyber incident in natural language and collects firmographic data, such as its country, industry category, age, annual revenue, and number of employees. The proposed model is integrated into the system to generate a larger Victim profile dataset and also to obtain firmographic data that users can use in the CRQ model. The full workflow of this model is illustrated in Figure 3.

- Retrieve basic information about the victim: using a cyber incident as a data source, a LLM is capable of extracting relevant information from the incident description such

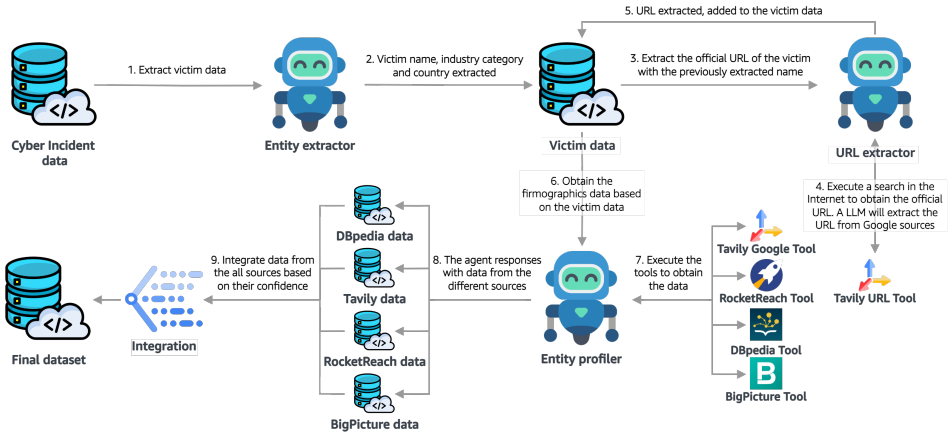

Fig. 3. Multi-agent architecture proposed for extracting victim information and completing missing firmographic data

as the name of the entity victim, the country and the industry. Additionally, the model queries various data sources from different websites on the Internet to retrieve the official website Uniform Resource Locator (URL). This basic information is used in subsequent processes to obtain or filter the firmographic data.

– Dig up firmographic information: with the basic information of the victim extracted, the model proceeds to perform queries across various external data sources to retrieve additional details. The information to be collected consists of eight distinct fields encompassing different data types: name of the victim, industry category, founded year, number of employees, country, revenue, publicly traded status, and North American Industry Classification System (NAICS) codes. The model requests information about the victim entity from four main sources: Google, DBpedia, BigPicture, and RocketReach. DBpedia and BigPicture are open-access sources available to any user seeking information on companies. In contrast, RocketReach is a paid platform that provides data with a higher level of reliability, as it is a more private and commercially maintained product. Google information is retrieved through a tool called Tavily, a platform that enables agent-based AI models to connect to the web. Tavily collects data from multiple Google sources while returning a filtered subset of relevant results and selecting only a reduced number of sources.
– Combine the extracted data: once the information has been extracted from the data sources, the final step performed by the model is the aggregation of the data to generate a consolidated set that captures the most complete information possible. To achieve this, a confidence score was assigned to each dataset retrieved from the sources based on the number of fields successfully extracted. The source with the highest confidence score was selected. In cases where certain firmographic fields are missing, the model attempts to replace them by merging data from other sources where their confidence scores exceed a defined threshold.

Once the final agent gathers all datasets from various sources, the goal is to consolidate them into a single dataset that maximizes data completeness. In particular, the final dataset

should include the victim's name, category, NAICS code, founded year, size, country, revenue, and public trading status. In addition to these fields, a confidence score was assigned to each source, represented as a numerical value between 0 and 1. This score reflects the reliability of the extracted information, completeness of the retrieved data, and number of relevant results obtained during the search process, rather than data quality. A value close to 0 indicates highly unreliable information, whereas a value of 1 indicates high reliability.

The value of the estimation is computed as shown in Equation 1 for each source s.

$$c\_fields_s = \frac{m_s}{f} \qquad c\_responses_s = \begin{cases} \frac{1}{n_s} & \text{if } n_s > 0 \\ 0 & \text{if } n_s = 0 \end{cases}$$
$$conf_s = \frac{1}{r_s} \cdot \Big[\alpha \cdot c\_fields_s + (1-\alpha) \cdot c\_responses_s\Big] \tag{1}$$

where the confidence score, $conf_s$, quantifies the reliability of a specific source. This confidence depends entirely on the completeness of the extracted data, the number of responses obtained, and the efficiency of the data retrieval process. The first component c_fields_s measures the number of fields obtained m_s relative to the total number of fields to be extracted f, that is, its proportion. The second component, $c\_responses_s$ serves as a confidence penalty when a source returns a large number of potential results, n_s, indicating uncertainty regarding the most accurate outcome; the confidence value is correspondingly reduced. The overall confidence score, $conf_s$, integrates these two components through a weighted average controlled by a tunable parameter, α, and is adjusted according to the type of source selected. For instance, in the case of $s =$ RocketReach a higher value is used to prevent the confidence from being penalized, as this source typically returns a large number of responses. The number of attempts, r_s, required to obtain a valid result increases with each failed query and always starts at one.

Queries are executed using a list of potential victims generated by the *Entity extractor* agent. This list contains extensive terminology referring to the victim, including variations such as corporations, brands, or subsidiaries, ordered from the most likely to the least likely, based on the LLM. Each agent sequentially queries the names until relevant information is obtained. Every failed attempt increments r, which reduces the final confidence score to reflect the inefficiency of the search. There is always an initial attempt; therefore, r_s can never be 0.

When the confidence $conf_s$ is computed for each source s, the dataset with the highest overall confidence is selected. If one or more fields are missing in the selected dataset, the system evaluates those specific fields across all sources with confidence scores higher than a threshold of 0.5 and aggregates the data.

Electronic Transactions Development Agency (ETDA) Threat Group Cards The ETDA is a Thai public organization that focuses on promoting secure and efficient digital transactions. Among its key initiatives, ETDA develops detailed Threat Group Cards to support CTI efforts. These profiles integrate data from leading global sources, including the Malware Information Sharing Platform (MISP) Threat Actors Galaxy, MITRE ATT&CK Framework, Malpedia, Open Threat Exchange (OTX), and ETDA's own CTI

archive and open-source research, offering a comprehensive view of threat actors targeting digital infrastructure.

Data acquisition was conducted through direct downloads of individual Threat Group Cards and a consolidated JavaScript Object Notation (JSON) file, both available on the official ETDA website. The consolidated file aggregates all documented actors and is designed for seamless integration with platforms such as MISP. Taken together, these two resources provide a rich set of attributes for each threat group, including their countries of operation, targeted industries, stated or inferred motivations, attributed campaigns and their timelines, as well as the tools and techniques employed.

Another important data source integrated into the system consists of cybersecurity vulnerabilities and related attack patterns, which are then used and completed by the CVE2TTs model (see Section 3.2). The process begins by collecting vulnerability data from the NVD, which includes CVE identifiers, descriptions, and associated CWE mappings. These CWEs are linked to CAPEC entries by MITRE, which in turn connect to MITRE ATT&CK techniques. While not every CVE has a direct mapping along this full chain, available mappings are used when possible, and gaps, particularly in the Industrial Control Systems (ICS) matrix, are predicted by the CVE2TTs model.

Internal telemetry data In addition to external sources, the system also incorporates internal data. However, this component is more limited, as obtaining this data from industrial organizations requires explicit authorization and faces significant operational and confidentiality constraints. Nevertheless, the system integrates internal telemetry whenever available, primarily consisting of anonymized samples of Intrusion Detection Systems (IDSs) data provided by DeNexus. It also includes detailed visibility into the number of connected assets, their vulnerabilities, and their distribution according to the Purdue model. This information is particularly valuable for risk quantification, as it enables the identification of exposed assets and weaknesses directly linked to the operational infrastructure. From this telemetry, the system extracts CVEs, which are then related to the CVE2TTs model (see Section 3.2). When such data are not accessible, users can alternatively upload a list of CVEs associated with the entity, ensuring that the quantification process remains consistent and applicable. By combining OT-specific IDS discovery capabilities with flexible input options, the platform strengthens both the accuracy and relevance of cyber risk assessment while maintaining confidentiality.

3.2. Indicators

To quantify cyber risk effectively, a set of indicators has been defined to capture distinct patterns and characteristics of the data, each serving a specific purpose depending on the indicator. Specially, in the DICYME project there are three main indicators: ATR2ATK, THRACT, and CVE2TTs.

ATR2ATK ATR2ATK or Attractiveness is an indicator for forecasting cyber incidents by assessing the proneness of an entity to them. It is decomposed into three main branches: basal attractiveness, online reputation, and potential victimisation.

Basal attractiveness focuses on inherent static characteristics, often referred to as firmographic data, such as location, operational criticality, data sensitivity, and size. These

attributes make certain entities more appealing targets for adversaries because of their intrinsic nature. It is modelled with association rule mining, specifically the FP-Growth algorithm, combined with a Decision Tree classifier that uses its output (amount of satisfied rules and aggregated support, lift and confidence per incident) [8, 7].

Online Reputation measures the presence of the entity in social networks and social media, as it may be more attractive to an adversary based on what users, customers, or employees write, communicate, and share anywhere on the Internet based on their perceptions and experience at any moment of their relationship, direct or indirect, with the entity [20, 21, 33]. It is a dynamic component and time-dependent, as it represents a specific moment, influenced by previous periods. Its computation relies on engagement (E), defined as the ratio of interaction (I) and reach (R) [3]. Reach indicates the number of individuals who view or mention a publication, whereas interaction (I) represents those who actively respond to it. In this context, engagement is assessed by considering whether the content originates from the entity itself through its official sources, where sentiment is assumed to be predominantly positive, or from other users, where additional sentiment analysis is taken into account since perceptions can be positive, neutral, or negative.

Lastly, potential victimisation, which is also a dynamic component, has to be with the visibility and technical knowledge that the adversary can have about the entity. Its rationale lies in the idea that organisations become more likely targets when they gain visibility in underground forums, leak sites, or other malicious communities, and when technical indicators reveal that their infrastructure may be easy to compromise. To capture this, two variables are monitored: the number of mentions in dark web leaks and the number of publicly visible assets connected to the Internet. These are combined through a buffer method that assigns a value between 0 and 2: 0 if neither variable is present, 1 if only one of them is, and 2 if both are detected. This provides a simple yet effective measure of potential victimisation.

THRACT This indicator is a composite metric designed to provide an overview of threat actor activity targeting a specific country and industry. Based on the data extracted from the ETDA, three partial metrics are derived to reflect their activity, capabilities, and objectives with respect to a specific victim. Activity is captured through the time elapsed since the actor was last observed. As the ETDA database is updated on a rolling basis, this value may change over time, reflecting updates in observed behavior. Actor capabilities are represented based on expert annotations of operational capacity. Victim-related features include country and industry, while actor motivation is also considered. Because the metric depends on the country and industry of the potential victim, it is dynamic rather than a fixed score for each actor in any situation, providing a context-sensitive measure of potential cyber risk.

CVE2TTs Finally, the CVE2TTs indicator corresponds to a ML model that maps CVE entries to specific Techniques (and, by extension, Tactics) within the MITRE ATT&CK framework. This mapping is performed by leveraging the textual description of each CVE together with additional attributes such as vulnerability type, CWE, CAPEC, and attack vector [29]. This concept is very important in the cybersecurity sector because it helps understand the practical exploitation of vulnerabilities in a structured manner, linking specific vulnerabilities to known adversary behaviors TTPs. This model is crucial for risk

management, CTI, and incident response, as it bridges the gap between vulnerabilities and the manner in which attackers might exploit them.

3.3. Cyber Risk Quantification (CRQ) model

In the context of cyber risk management, the Cyber Risk Quantification model introduced in this proposal enables a simulation-based tool that quantifies cyber risk in financial terms, supporting informed decision-making for organizations. Using stochastic simulations, the model estimates the Annual Expected Loss for a given organization by combining the frequency of successful attacks and their financial magnitude or impact, as shown in Equation 2.

$$\text{Cyber risk} = \text{Loss Event Frequency (LEF)} \times \text{Loss Magnitude (LM)} \tag{2}$$

The model relies on a state-of-the-art risk tree that separates probability and impact into two main branches. The first component, the Loss Event Frequency (LEF), represents the expected number of loss events within a given period, typically one year. In this context, loss events correspond to materialized incidents that cause harm to the organization. The LEF is determined as the product of the Threat Event Frequency (TEF), which estimates the number of attempted attacks that the organization is likely to face, and the Susceptibility, which reflects the probability that those attempts will successfully translate into incidents. This relationship can be expressed as shown in Equation 3.

$$\text{Loss Event Frequency (LEF)} = \text{Threat Event Frequency (TEF)} \times \text{Susceptibility} \tag{3}$$

This formulation integrates both the external pressure from the threat landscape (captured by the TEF), and the internal defensive posture of the organization (captured by Susceptibility). The TEF takes into account a baseline number of incidents based on private cyber intelligence knowledge, the ATR2ATK indicator, and the annualized rate of incidents for the country and industry category of the entity, sourced from the Cyber incident dataset. Susceptibility is derived from the vulnerabilities obtained through internal telemetry or from a list of vulnerabilities uploaded by the user, in combination with the CVE2TTs model, the Threat Actor Index, which captures the activity and focus of the 245 actors in the database over a three-year period, and an adjustable security profile index. The latter characterizes the defensive posture of the organization by incorporating features such as organizational size, number of unpatched vulnerabilities, and exposure of devices visible from the internet.

The second component of the risk tree is the Loss Magnitude (LM), which quantifies the financial impact of a successful attack as the combination of primary and secondary losses, as stated in Equation 4. The identification of feasible primary and secondary losses depends on the attack technique k, which is constrained by the vulnerabilities observed or uploaded by the user, linked to the CVE2TTs mapping. In practice, just the achievable techniques from the Impact MITRE ATT&CK tactic, given the vulnerability landscape of the entity under analysis, are considered in the simulation. This ensures that both primary and secondary losses are not only probabilistically derived, but also grounded in the

technical conditions that define the exposure of the organization. Expert financial knowledge is further incorporated to parametrize and calibrate the estimation of losses to reflect realistic economic consequences for the entity.

$$\text{Loss Magnitude (LM)}_k = \text{Primary Loss}_k + \sum \text{Secondary Losses}_k \tag{4}$$

Primary losses represent the direct effect of an incident and include categories such as business interruption, equipment damage, extortion, or human impact. In each simulation run, only one primary loss type is selected, corresponding to the most plausible impact mechanism for the attack technique under consideration. By contrast, secondary losses capture indirect consequences, such as forensic investigation costs, reputational damage, or regulatory penalties. Unlike primary losses, several secondary losses can occur simultaneously, and their combined effect is represented by summing the estimated financial amounts for each of them.

3.4. Visualization and decision support

The visualization and decision support component is implemented as an interactive application that allows users to explore datasets and execute risk models through two main options: either using the entities defined in the Victim profile dataset or by providing their own model inputs. The platform provides a wide range of interactive visualizations, including bar charts, histograms, stacked and unstacked time series, heatmaps, and world maps. Figure 4 shows some of this techniques used inside the CVE2TTs module, reflecting the relations made between CVEs and MITRE ATT&CK Tactics and Techniques by vulnerability type and by Common Vulnerability Scoring System (CVSS), as well as the number of vulnerabilities and predicted techniques per year. Users can filter the data by relevant variables such as dates, incident databases, CVE types, and other context-specific metrics. This interactivity enables stakeholders to examine patterns in risk indicators, compare scenarios, and gain a deeper understanding of the underlying data and model outputs.

In addition, the system generates a downloadable report that compiles all the data, the computed indicators, and details on the steps of the quantification process. This report can be shared with other stakeholders, including expert panels, managers, or decision-makers, ensuring that risk analyses can be reviewed, discussed, and extended beyond the interactive platform. By combining interactive exploration with exportable documentation, the system supports both immediate, hands-on decision-making and broader strategic planning across multiple levels of an organization.

3.5. Recommender model

In addition, the system incorporates a set of recommender models that complement its core functionality. These models are designed to identify patterns within the data and indicators, extract meaningful features, and provide tailored recommendations. Although is not essential for the system baseline operation, this component enhances its analytical capabilities and offers added value by supporting more informed decision-making. The models employed to perform these tasks were fine-tuned through prompt engineering [11] and Retrieval Augmented Generation (RAG) techniques [15] to produce responses

Fig. 4. Example of visualization techniques used in the CVE2TTs module of the DICYME system. The top heatmap displays the distribution of CVEs identified for each ICS tactic by vulnerability type. The middle bar chart illustrates the temporal evolution of security data, showing the total annual count of CVEs (blue) alongside the volume of predicted Enterprise-specific techniques (orange) and ICS-specific techniques (yellow) associated with those vulnerabilities. Finally, the bottom pyramid chart presents the average CVSS scores per tactic, calculated from the CVEs associated with each tactic according to the CVE2TTs model's predictions. The chart is bifurcated by domain: the Enterprise matrix tactics are positioned on the left, while ICS matrix tactics are on the right. Within this chart, blue bars indicate average scores according to the CVSS v2 standard, and orange bars represent the CVSS v3 scoring

that are as accurate as possible. By adapting the prompts for each system component with their corresponding data, a specialized prompt is generated for that specific component to ensure a more accurate and context-aware response.

In some cases, the data used in the prompt processed by the model may lead to undesired responses. This occurs because the system generates large volumes of data that may pose challenges for the model when generating an analysis, since these models are subject to token limitations. To address this issue, query decomposition [36] is applied in situations where sections contain extensive data. The queries were divided, allowing multiple interactions with the model to provide more specific explanations. This approach made it possible to explain similar datasets without resorting to generalization. Finally,

the responses to these queries were consolidated by the LLM using a prompt designed to summarize all partial outputs and deliver a comprehensive final explanation.

In addition to the models that provide data explainability, decision support capabilities, and parameter adjustment, the system also incorporates a more sophisticated model implemented as a chatbot assistant. The assistant enhances user interactivity by enabling the execution of the same tasks performed by other models with the added flexibility of allowing users to request specific adjustments or obtain tailored explanations through natural language queries. The aim of enhancing user interaction, explainability, and adjustments is achieved through the assistant, which is integrated into one of the most important and critical components of the system: the CRQ simulation. Through an intuitive interface, users can submit questions related to the CRQ, which the AI assistant answers one at a time in a conversational manner. The model has a limited memory capacity, allowing it to retain awareness of previous prompts and generate responses within a specific window. Thanks to this memory, the assistant can access prior questions and answers, compare its outputs, retrieve the results of previously executed simulations, and revisit earlier recommendations. This enables the contrast of past outcomes with newly generated ones, thereby supporting more consistent analyses and more informed decision-making.

The architecture implemented for the GenAI models is inspired by the approach proposed by DeepSeek-V3 [14], which is based on a Mixture of Experts (MoE) model. This architecture uses a set of specialized experts, each responsible for handling specific tasks depending on the context. The assistant is capable of handling three distinct tasks, each managed by a dedicated agent within the architecture: answering general questions related to CRQ process, providing recommendations on parameter adjustments, and rerunning the simulation to generate a comparison with the previous simulation results.

4. Experiments

The experiments presented in this study evaluated two key aspects of the proposed system. First, we assess the data extraction and integration capabilities of the multi-agent architecture, focusing on the completeness and quality of the firmographic information collected from multiple cyber-incident databases.

Second, we evaluated the entire risk quantification workflow, from the computation of different intermediate indicators and metrics present in the system to the computation of risks, probabilities, frequencies, and other risk parameters. For this purpose, a simulation will be conducted using the model with a case study and real data from a possible victim entity or institution.

The model selected for the multi-agent architecture, which extracts the victim data, is the Meta LLaMa 4 Scout [18], with a total of 17B active parameters and the advantage of being open-weight, which allows us to maximize utility. This model was chosen over other available alternatives because, in addition to its ease of deployment, it supports the execution of agents and tools through libraries employed in this project.

4.1. Data gathering

Within the DICYME application, a wide range of datasets are employed across different tabs to present information using tables, charts, and other visualization components. As

previously discussed, the data were extracted from various sources and organized into three main datasets used in different modules of the application: Cyber incidents, Victim profile, and IDSs.

This experiment evaluates the firmographic data completion of the VISCA module by comparing the dataset it generates against the analyst-annotated Victim profile dataset. After collecting information from victim entities affected by cyber incidents, the data are stored in a dataset to be used within the system. The objective of this experiment was to compare the automatically extracted dataset with a similar dataset in which the victim entity data were manually collected by a human annotator.

Conversely, the victim dataset used to evaluate the extracted dataset was a fully handcrafted dataset compiled by a human annotator. It contains 675 instances of distinct victim entities and includes a substantial amount of information related to each company, specifically firmographic data. In particular, the dataset contains 9 variables, including the incident category, entity, industry category, country, earnings, employees, revenue, profitable and publicly traded. This dataset is used within the system to compute various indicators and represent key information about each victim entity, which explains the large number of variables included, as they serve as inputs for these indicators.

In contrast, the dataset under evaluation, obtained after the extraction and aggregation of firmographic data, was automatically constructed by a multi-agent system capable of executing a workflow starting from a cyber incident. Once the agent extracts data from multiple data sources, it integrates them into a final output to produce a consolidated result. The final output contains eight firmographic fields, including: entity, category, country, revenue, founded year, employees, publicly traded, and NAICS code. This dataset contains 1,658 observations of distinct victim entities processed from cyber incidents collected from the ICS STRIVE, EuRepoC, and TI Safe databases. In all the processed cyber incidents, at least one victim entity was explicitly identified in the incident description, which enabled the extraction of the corresponding firmographic data.

Table 1 presents a comparison between the two datasets used to model a profile for a victim entity. It can be observed that the VISCA-extracted dataset contains a significantly larger number of identified victim entities. This is one of the most notable characteristics of the model, as its autonomous workflow can continue processing incidents until all available incidents are processed, while also achieving a much higher processing speed than manual annotation can. This difference in gathering speed is evident in Table 1, which includes the range of extraction dates for the data collected. While the Victim profile dataset required an entire year to compile all the victim entities it contained, the VISCA model was able to extract more than double the number of victim entity records in just five months of uninterrupted execution.

However, one limitation of this dataset is the smaller number of variables compared to the Victim profile dataset. This difference is due to the specific original goal of the VISCA model: to generate victim-related data for cyber risk quantification using the CRQ framework. In contrast, the Victim profile dataset was developed to support multiple components of the system, including the computation of indicators, metric computation, and structured information representation. Nevertheless, the VISCA model can be extended to collect additional firmographic attributes following the same automated process, allowing the resulting dataset to become more comprehensive and usable across a wider range of system components.

Table 1. Comparison between the Victim profile dataset and the VISCA-extracted dataset

Feature / Metric	Victim profile dataset	VISCA-generated dataset
Unique entities	675	1,658
Number of fields	9	8
Sources used	Manual collection	ICS STRIVE, EuRepoC, TI Safe
Extraction method	Manual annotation by an analyst	Multi-agent automated workflow
Extraction dates	December 2023 - December 2024	May 2025 - September 2025

4.2. Cyber Risk Quantification (CRQ)

To evaluate the effectiveness of the CRQ procedure, we conducted a series of experiments that simulate real-world scenarios using actual data from multiple companies. These experiments were designed with two main objectives: first, to demonstrate the internal performance of the CRQ process, including the types of input data and datasets that can be leveraged within this framework, and second, to validate the output risk metrics generated by the model, which serve as a foundation for informed decision-making by organizations.

Executing the CRQ workflow of the system requires the input of various firmographic attributes that characterize a company. On the one hand, firmographic data from real companies can be loaded from the Victim profile dataset, specifically including country, industry category, revenue, earnings, number of employees, publicly traded status, profitability, online reputation, publicly visible devices, and critical information leaks. However, these fields can be populated directly with the data of the company provided by the user operating the system. In addition, other simulation parameters are configurable, such as the random seed, number of simulations, and insurance parameters. Finally, a list of unpatched CVEs present in the infrastructure of the organization, which may be exploited, must be provided, as they serve as a basis for estimating the corresponding risk measures, as stated in Section 3.3.

For the experiments, a real-world case was selected, corresponding to an actual company extracted from our dataset. The input values for the model are the previously described firmographic attributes of the company, all of which are available in the dataset. Specifically, the University of Michigan was selected for this experiment. This public institution was selected because it is recognized as one of the leading universities worldwide and is located in the United States, a country that consistently reports the highest concentration of cyberattacks. This combination of global academic relevance and geographical context makes the institution particularly susceptible to being frequently targeted by threat actors.

This public institution, located in the United States and dedicated to education, reported a revenue of €9.1 billion, employed a total of 34,600 staff members, and had three publicly visible devices and one critical information leak. Using these publicly available data from the university, CRQ simulations were executed and complemented with additional configuration parameters, including the number of runs, random seed, and insurance settings. The CRQ procedure first computes a series of indicators that are subsequently used to derive the output of the model parameters. The values of all the indicators are listed in Table 2. Among them, the attractiveness score is 80%, a significantly high value, primarily driven by the institution's location in the United States and its educational industry category. Another key indicator is the Threat Actor Index, which reached a value

Table 2. Summary of the values obtained from the CRQ simulation, including both the computed indicators and the model output parameters

Indicators	Baseline	85 events/year
	Incident rate	50%
	Attractiveness	80%
	Threat actor index	58%
	Exposure	41.61
	Security profile	67%
CRQ Outputs	Threat Event Frequency (TEF)	34 events/year
	Susceptibility	2%
	Loss Event Frequency (LEF)	15 events/year
	Primary loss	€49,512
	Secondary loss	€20,786,315
	Loss Magnitude (LM)	€20,835,828/event
	Cyber risk	€14,168,363/year

of 58%, reflecting the intensity and capabilities of threat actors over a three-year period and corresponding to a moderate risk level.

Based on the computed values derived from the publicly available data of the university and complemented with additional simulation parameters, such as the number of runs, a random seed to ensure reproducibility, and cyber insurance financial values, the model produces simulation outputs presented in both graphical and numerical forms. Figure 5 presents the results of the simulated risk through two charts: the first describes the distribution of loss events across all simulated years (each simulation means a year), while the second illustrates the Loss Exceedance Curve (LEC), which illustrates the probability (Y-axis) that cyber risk losses will be equal to or greater than a given amount (X-axis).

In the first chart, which represents the distribution of annual loss events, the results show a left-biased pattern. The majority of simulated years (574 out of 1,000) resulted in no materialized incidents, while 306 years included a single loss event. Multiple events within the same year were comparatively rare: 83 simulations produced two events, 24 produced three, 9 produced four, and only 4 reached five events. This highlights that, although the occurrence of cyber loss events is possible, their frequency per year remains predominantly low, consistent with the heavy-tailed nature of cyber risk distributions.

The second chart, which illustrates the distribution of financial losses (LEC), reveals two distinct components. At the lower end, there is a steep initial drop, corresponding to rare but high-impact scenarios—so-called black swan events. Beyond this abrupt decline, the curve follows an almost linear descent, decreasing steadily from a 47% probability of incurring losses equal to or greater than €8.9×10^5 down to 0% probability at €9.0×10^7. Unlike sharply convex loss distributions, this profile suggests a smoother, more gradual reduction in probability across a wide range of loss magnitudes, balancing rare catastrophic scenarios with a persistent probability of moderate to high losses.

Finally, the CRQ model also provides single averaged values, complementing the previously described distributions and probabilities, to facilitate a direct quantification of losses and frequencies. Based on the indicators and other adjustable system parameters, the CRQ framework estimates a LEF of 0.68 events per year and a LM of €20,835,828 per event. By combining these two measures, the system produces the annualized risk

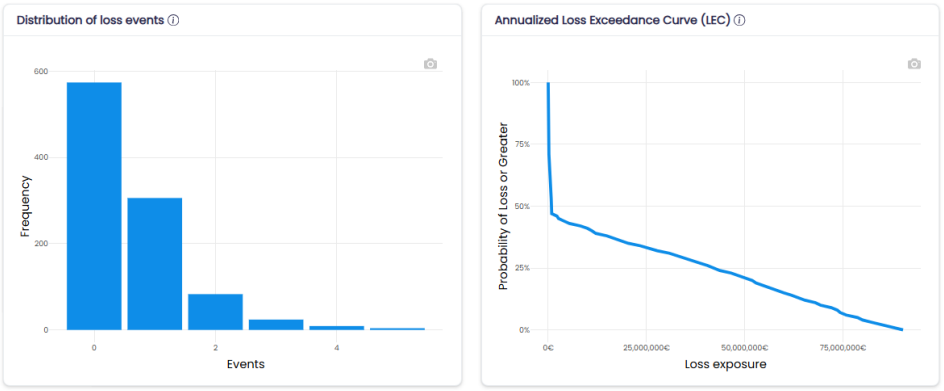

Fig. 5. Distribution of Simulated Loss Events and Corresponding Loss Exceedance Curve in the CRQ simulation

value for the organization, which in the case of the University of Michigan is estimated at €14,168,363 per year.

5. Lessons learned

The analysis of the experimental results provides several insights into the strengths and limitations of this proposed system. The experiments conducted on the CRQ model revealed both notable advantages and certain weaknesses of this component. One of these strengths is the flexibility to simulate realistic scenarios using firmographic data from real companies, while also allowing the introduction of fictitious parameters that emulate organizations not present in the dataset. This strength is further enhanced by the ability to leverage firmographic information obtained from the Victim profile and VISCA datasets, which enrich the inputs of the model and provide a stronger foundation for risk quantification. This capability greatly expands the range of situations in which the model can be applied, as it not only quantifies risks for entities included in the repositories but also provides valuable estimations for hypothetical cases, making it useful in decision-making and what-if analyses. Additionally, the results generated by the model are highly descriptive and presented through graphical visualizations that allow users to easily interpret the represented case. By combining frequencies, probabilities, distributions, loss estimations, and a final risk value, the model provides a comprehensive and visually intuitive quantification of the different possible risk situations.

However, a key limitation of the current evaluation is the absence of formal validation of the accuracy of the quantification and recommendation strategies. A potential approach to address this limitation is post-mitigation validation, leveraging both publicly available but also internal data from organizations that have experienced cyberattacks. This would enable comparisons between the predicted frequencies and loss estimations by the model against the actual incidents and financial losses reported over a given period, providing an empirical basis for assessing the reliability of the CRQ outputs. In the best-case scenario, although it may not be scalable, organizations could grant controlled access to private

datasets containing historical records of cyberattacks, including incidents that occurred before and after the implementation of specific defensive measures. This would allow the system to assess whether the predicted risk metrics accurately reflect the real-time impact of mitigation strategies.

The experiments also revealed that many records in the evaluated datasets contain missing fields, either because the information could not be found or because it simply does not exist. When key variables are absent, the model cannot perform optimal estimations. This issue was evident, for example, in the Victim profile dataset, where the used data sources almost never provide the earnings field. The absence of inputs may lead to incorrect probability distributions or models, having to consider whether it can be used or it is better to exclude them from the analysis.

Finally, the experiments conducted on the data extraction model demonstrated the speed and efficiency of aggregating victim entity information in a concrete case of cyber-incident processing. This capability highlights the potential of the system to automate time-consuming tasks, such as data labelling, enabling the generation of a fully usable dataset for the system. In addition to automating the process, the model operates significantly faster than manual human processing, allowing the creation of large-scale datasets that would otherwise require substantial time by humans. These enriched datasets can then be leveraged by other components of the framework, such as the CRQ model, thereby enhancing their performance by incorporating more relevant and updated firmographic information about organizations. However, an important limitation of the current approach is that the extracted data are not validated; the system only checks if the retrieved fields contain a value, but not whether the value is factually correct. Although the workflow relies on trusted and updated data sources, it also incorporates information retrieved from open Internet searches, which introduces the risk of inaccurate or inconsistent data due to the large number of possible sources.

6. Conclusion

This study introduced DICYME, a complex system that contains multiple datasets covering different cyber-related domains, information representation and processing capabilities, and various techniques designed to identify patterns and characteristics within the data. Throughout this study, the methods and techniques used to implement the complex workflow of the system are presented, beginning with the acquisition of diverse datasets that are subsequently employed to compute a range of indicators for quantifying cyber risk across all instances. Subsequently, the CRQ model was proposed, enabling a large number of simulations aimed at quantifying the previously collected data and their computed indicators, and finally producing a final risk value along with other relevant metrics.

In the experiments section, the results demonstrated that both useful components of the system provide significant capabilities: a comparison of the Victim profile dataset and the CRQ model. The VISCA model exhibited high performance in the data extraction task, achieving efficient execution times and demonstrating the potential to fully automate the annotation task performed by a human. However, an important limitation of the current approach is that during data extraction, the system does not explicitly obtain the most accurate value among all possible sources. This limitation may lead to subsequent components and models of the system with inaccurate or inconsistent data, which can

propagate errors into subsequent stages of the system. As future work, it would be important to implement a mechanism that not only maximizes the number of extracted attributes but also ensures the selection of the most reliable value for each field. Additionally, extending the set of collected variables would enable a more detailed and representative modelling of the victim profiles.

Similarly, the CRQ model proved capable of generating a rich set of metrics and quantitative outputs, offering valuable insights that can directly support post-mitigation actions by helping organizations prioritize defensive strategies. Nevertheless, there is currently no validation to confirm whether the estimations of the model accurately reflect real-world outcomes. This limitation is critical, as both the overestimation of risks and the underestimation of losses or attack frequencies could lead organizations to make incorrect decisions. Future work should include rigorous validation of the CRQ outputs, either by leveraging public datasets from organizations that have experienced cyberattacks or through partnerships with private companies to provide access to historical incident data. Such validation would allow for a systematic comparison between the predicted annual losses and event frequencies by the model and actual observed outcomes.

Acknowledgments. This work has been funded by the Spanish MICIU under the CPP program in the DICYME project (Ref: CPP2021-009025), partially funded by the XMIDAS project (PID2021-122640OB-I00), and supported by DeNexus Inc.

References

1. Center for International and Security Studies at Maryland: Cyber Events Database (2024), https://cissm.umd.edu/cyber-events-database
2. Chu, Z., Wan, Y., Li, Q., Wu, Y., Zhang, H., Sui, Y., Xu, G., Jin, H.: Graph neural networks for vulnerability detection: A counterfactual explanation. In: Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis. pp. 389–401 (2024)
3. Cioppi, M., Curina, I., Forlani, F., Pencarelli, T.: Online presence, visibility and reputation: a systematic literature review in management studies. Journal of Research in Interactive Marketing 13(4), 547–577 (2019)
4. Electronic Transactions Development Agency: Threat Group Cards: A Threat Actor Encyclopedia (2025), https://apt.etda.or.th/cgi-bin/aptgroups.cgi
5. EuRepoC project: The European Repository of Cyber Incidents (2025), https://eurepoc.eu/
6. Falowo, O.I., Popoola, S., Riep, J., Adewopo, V.A., Koch, J.: Threat actors' tenacity to disrupt: Examination of major cybersecurity incidents. IEEE access 10, 134038–134051 (2022)
7. García-Ochoa, J., Fernández-Isabel, A., Contreras, C., Fernández, R.R., de Diego, I.M., Beltrán, M.: Defining the attractiveness concept for cyber incidents forecasting. Computer Science and Information Systems (ComSIS) (2025)
8. García-Ochoa, J., Fernández-Isabel, A., Martín de Diego, I., Contreras, C., R. Ravines, R., López, O.: Defining the basal attractiveness concept for cybercriminals. In: Proceedings of the 10th Jornadas Nacionales de Investigación en Ciberseguridad. pp. 513–517. Universidad de Zaragoza, Zaragoza, Spain (2025)
9. Industrial Safety and Security Source and Waterfall Security Solutions: ICSSTRIVE (2025), https://icsstrive.com/
10. Kamruzzaman, M., Bhuyan, M.K., Hasan, R., Farabi, S.F., Nilima, S.I., Hossain, M.A.: Exploring the landscape: A systematic review of artificial intelligence techniques in cybersecurity. In: 2024 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI). pp. 01–06. IEEE (2024)

11. Knoth, N., Tolzin, A., Janson, A., Leimeister, J.M.: Ai literacy and its implications for prompt engineering strategies. Computers and Education: Artificial Intelligence 6, 100225 (2024)
12. KonBriefing: Cyberattacks, Hacker attacks, Ransomware attacks (2025), https://konbriefing.com/en-topics/cyberattacks.html
13. Landauer, M., Skopik, F., Stojanović, B., Flatscher, A., Ullrich, T.: A review of time-series analysis for cyber security analytics: from intrusion detection to attack prediction. International Journal of Information Security 24(1), 3 (2025)
14. Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., et al.: Deepseek-v3 technical report. arXiv preprint arXiv:2412.19437 (2024)
15. Liu, F., Kang, Z., Han, X.: Optimizing rag techniques for automotive industry pdf chatbots: A case study with locally deployed ollama modelsoptimizing rag techniques based on locally deployed ollama modelsa case study with locally deployed ollama models. In: Proceedings of the 2024 3rd International Conference on Artificial Intelligence and Intelligent Information Processing. pp. 152–159 (2024)
16. Liu, R., Xie, Y., Dang, Z., Hao, J., Quan, X., Xiao, Y., Peng, C.: Dynamic vulnerability knowledge graph construction via multi-source data fusion and large language model reasoning. Electronics 14(12), 2334 (2025)
17. Mallick, M.A.I., Nath, R.: Navigating the cyber security landscape: A comprehensive review of cyber-attacks, emerging trends, and recent developments. World Scientific News 190(1), 1–69 (2024)
18. Meta: Introduction to the LLaMA 4 Models (2025), https://www.llama.com/docs/model-cards-and-prompt-formats/llama4/
19. Mızrak, F.: Integrating cybersecurity risk management into strategic management: a comprehensive literature review. Research Journal of Business and Management 10(3), 98–108 (2023)
20. Okutan, A., Werner, G., Yang, S.J., McConky, K.: Forecasting cyberattacks with incomplete, imbalanced, and insignificant data. Cybersecurity 1, 1–16 (2018)
21. Okutan, A., Yang, S.J., McConky, K., Werner, G.: Capture: cyberattack forecasting using non-stationary features with time lags. In: 2019 IEEE Conference on Communications and Network Security (CNS). pp. 205–213. IEEE (2019)
22. Paolo Passeri: 2024 Cyber Attacks Statistics (03 2024), https://www.hackmageddon.com/2024/03/26/2024-cyber-attacks-statistics/
23. Paolo Passeri: Hackmageddon (2025), https://www.hackmageddon.com/
24. Paté-Cornell, M.E., Kuypers, M., Smith, M., Keller, P.: Cyber risk management for critical infrastructure: a risk analysis model and three case studies. Risk Analysis 38(2), 226–241 (2018)
25. Phillips, S.C., Taylor, S., Boniface, M., Modafferi, S., Surridge, M.: Automated knowledge-based cybersecurity risk assessment of cyber-physical systems. IEEE Access 12, 82482–82505 (2024)
26. Pseftelis, T., Chondrokoukis, G.: Understanding cyber incident dynamics in the european union: A study of actor types and sector vulnerabilities. Preprints.org (04 2025)
27. Qin, X., Jiang, F., Cen, M., Doss, R.: Hybrid cyber defense strategies using honey-x: A survey. Computer Networks 230, 109776 (2023)
28. Qudus, L.: Advancing cybersecurity: strategies for mitigating threats in evolving digital and iot ecosystems. Int Res J Mod Eng Technol Sci 7(1), 3185 (2025)
29. Rajesh, P., Alam, M., Tahernezhadi, M., Monika, A., Chanakya, G.: Analysis of cyber threat detection and emulation using mitre attack framework. In: 2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA). pp. 4–12. IEEE (2022)
30. Rios Insua, D., Couce-Vieira, A., Rubio, J.A., Pieters, W., Labunets, K., G. Rasines, D.: An adversarial risk analysis framework for cybersecurity. Risk Analysis 41(1), 16–36 (2021)
31. Sailio, M., Latvala, O.M., Szanto, A.: Cyber threat actors for the factory of the future. Applied Sciences 10(12), 4334 (2020)
32. Shi, Z., Matyunin, N., Graffi, K., Starobinski, D.: Uncovering cwe-cve-cpe relations with threat knowledge graphs. ACM Transactions on Privacy and Security 27(1), 1–26 (2024)

33. Subroto, A., Apriyana, A.: Cyber risk prediction through social media big data analytics and statistical machine learning. Journal of Big Data 6(1), 50 (2019)
34. TI Safe: Incident Hub (2024), https://hub.tisafe.com/
35. Wåreus, E., Hell, M.: Automated cpe labeling of cve summaries with machine learning. In: International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment. pp. 3–22. Springer (2020)
36. Xie, Y., Jin, X., Xie, T., Lin, M., Chen, L., Yu, C., Cheng, L., Zhuo, C., Hu, B., Li, Z.: Decomposition for enhancing attention: Improving llm-based text-to-sql through workflow paradigm. arXiv preprint arXiv:2402.10671 (2024)
37. Yin, J., Hong, W., Wang, H., Cao, J., Miao, Y., Zhang, Y.: A compact vulnerability knowledge graph for risk assessment. ACM Transactions on Knowledge Discovery from Data 18(8), 1–17 (2024)
38. Zhou, Y., Liu, S., Siow, J., Du, X., Liu, Y.: Devign: Effective vulnerability identification by learning comprehensive program semantics via graph neural networks. Advances in neural information processing systems 32 (2019)

Javier Sánchez García-Ochoa was born in Toledo, Spain in 2000. He received a bachelor's degree in Cybersecurity Engineering from Rey Juan Carlos University (URJC) and a master's degree in Cybersecurity and Privacy from the Open University of Catalonia (UOC). After gaining more than a year of experience in the private sector, he joined Rey Juan Carlos University in 2023 as a research staff member. During this period, he participated in the public-private collaboration project DICYME (Dynamic Industrial Cyber Risk Modelling based on Evidence), where he focused on dynamic cyber risk quantification. He has contributed to several scientific articles and conference proceedings. His research interests currently focus on large language models (LLMs), data science, machine learning, and their applications within the field of cybersecurity.

Jaime Rueda Carpintero was born in Madrid, Spain in 2002. He received a degree in Computer Engineering from Rey Juan Carlos University (URJC), where he also completed a Master's degree in Decision Systems Engineering. After more than a year of experience in the private sector, he joined Rey Juan Carlos University in 2025 as a researcher in the DICYME project. He is a member of the High Performance Research Group Data Science Laboratory (DSLAB, URJC) and has contributed to several scientific articles and conference communications. He actively participates in research tasks and projects focused on natural language processing (NLP), large language models (LLMs), and data science.

Rubén Rodríguez Fernández was born in Bembibre, León, Spain in 1993. He received a PhD degree in Artificial Intelligence from Rey Juan Carlos University (URJC). He also received a master's degree in Data Science from URJC and a master's degree in Artificial Intelligence from the Technical University of Madrid (UPM). He is part of the Data Science Laboratory high performance research group and has been an Assistant Professor at URJC since 2024. His research interests include active learning, explainable machine learning, generative artificial intelligence, and applied machine learning.

Alberto Fernández-Isabel was born in Toledo, Spain in 1984. He received a PhD in Computer Science from Complutense University of Madrid (UCM) in 2015. He obtained

a scholarship at the Spanish National Research Council (CSIC) as a technical assistant. He has been working for several years on European and national projects as a predoctoral and postdoctoral researcher. Since 2019 he is Assistant Professor at the Higher Technical School of Computer Engineering (ETSII) at Rey Juan Carlos University (URJC). He has authored more than 30 scientific articles and books. He completes his background with a Master´s degree in Artificial Intelligence and a Master´s degree in Information Systems. His research interests include intelligent agents, machine learning, data visualization, and natural language processing in various application domains, including distributed programming, sentiment analysis, agent-based collaboration and negotiation, smart cities, and simulations

Isaac Martín de Diego was born in Campaspero, Valladolid, Spain in 1973. He received a PhD degree in Mathematical Engineering from Carlos III de Madrid University in 2005 (Extraordinary Doctorate Award). Since 2023 he is a full professor at the Higher Technical School of Computer Engineering at Rey Juan Carlos University (Associate Professor from 2018). He is the co-founder of the Data Science Laboratory and Head of the Sports Analytics Master at Rey Juan Carlos University. He has been head of the ERICSSON Chair on Data Science applied to 5G. He is the author of more than 100 articles. His research interests include methods, processes, and tools for Data Science in various application domains: explainability, sampling, complexity, performance evaluation, visualization, recommendation systems and security with a special interest in Machine Learning algorithms and a combination of information methods.

Emilio López Cano was born in Madrid, Spain in 1973. He received a PhD degree from Rey Juan Carlos University (URJC). He holds a Master's degree in Decision Systems Engineering (URJC), a degree in Applied Statistics, and a diploma in Statistics from the Complutense University of Madrid (UCM). He is an Associate Professor at Rey Juan Carlos University in the area of Statistics and Operations Research. Previously, he was a teacher and researcher at the University of Castilla-La Mancha since 2011, and worked as a statistician and IT professional in the private sector for 14 years. His research interests include statistical computing, analysis, visualization, and management of complex data and modeling in multidisciplinary environments, where he has several national and international publications and projects.

Romy Rodríguez Ravines received a PhD degree in Statistics from the Federal University of Rio de Janeiro (UFRJ), Brazil. She is an Applied Data Science Leader with over 25 years of experience in business consulting based on advanced analytics, Bayesian modeling, and artificial intelligence, specializing in transforming data into strategic value across the cybersecurity, finance, marketing, and public health sectors. She has served as the Head of Research and Modeling Strategies at DeNexus, where she led the development of a cyber risk quantification platform adopted by international insurers and industrial organizations. She has held various senior leadership positions in consulting firms and has served as a lecturer at several leading universities and business schools in Spain.

Ovidio López, born in Murcia, Spain, holds a Physics degree from the University of Murcia and a Ph.D. in Renewable Energy and Energy Efficiency from the Polytechnic University of Cartagena. His background includes numerical simulation, data analysis,

and Operational Technology (SCADA, PLC). He has experience in R&D and Machine Learning, later working as a Data Scientist focused on industrial cyber risk. Currently, he teaches Electronics in vocational training, sharing expertise in automation and emerging technologies.

Jaume Puigbò Sanvisens was born in Barcelona, Spain. He holds a Bachelor's degree in Mathematics from the University of Barcelona, as well as a Master's degree in Foundations of Data Science (2016–2017) and a Postgraduate degree in Data Science (2015–2016) from the same institution. Since 2021, he has been working as a Data Scientist at DeNexus Inc., where he develops advanced models to quantify cyberattack risk in industrial environments. His research and professional interests include Machine Learning, Deep Learning, Optimization, Recommender Systems, Natural Language Processing, and Data Visualization, with applications in cybersecurity, predictive analytics, and industrial risk modeling. In the past, he has worked on projects such as predictive maintenance for industrial systems, demand forecasting, traffic prediction, and B2B recommender systems, applying advanced statistical and machine learning techniques to real-world challenges. Jaume was a finalist in the UniversityHack 2018 competition and has presented at R user conferences.

Received: October 30, 2025; Accepted: Devember 22, 2025.