

**A DATA-DRIVEN THREAT MODELLING LANGUAGE (d-TM)  
FOR  
ENSURING CYBER SECURITY ASSURANCE**

**Mohammed K. S. Alwaheidi**

A thesis submitted in partial fulfilment of the requirements of the University of East London for  
the degree of Doctor of Philosophy

October 2023

## **Dedication**

In dedicating this thesis, I want to express my appreciation for the pillars of my life. I am profoundly grateful to my parents, whose unwavering faith, sacrifices, and guidance have always been a guiding light for me in the values of dedication and commitment. I also want to extend my heartfelt thanks to my wonderful wife; she has been a source of support, understanding and love throughout this challenging journey. Last but not least, I want to acknowledge my children surrounding me as they continue to inspire and motivate me to work hard and move forward. This work serves as a testament to the strength, love, and encouragement that all of you have given to me.

## ABSTRACT

In the context of a rapidly evolving business environment characterized by persistent cyber threats targeting enterprises and the introduction of new attack vectors through technological advancements, the significance of data intelligence is experiencing exponential growth, and ensuring uninterrupted business operations becomes increasingly challenging. To effectively defend against these threats and gain a comprehensive understanding of security posture, organizations must evaluate their digital infrastructure. Threat modelling is essential for understanding system threats, mitigating risks from current weaknesses, and developing strategic countermeasures for improving cybersecurity posture. Threat modelling systematically analysing the complex relationship between digital infrastructures, applications, and potential attackers. Threat modelling is a challenging task due to the variety of generated, stored, or processed data by digital infrastructure. However, many existing methodologies for threat modelling often struggle to examine and prioritize data-related threats. This highlights the need for innovations in this field to guarantee comprehensive security assurance. The research methodology encompasses Four stages, literature review, then model and tool development, lastly the evaluation and conclusion. Each part contributes to the overall development and evaluation of the data-driven threat modelling and analysis approach.

This thesis contributes a novel threat modelling approach to address the aforementioned challenges. The proposed model, known as d-TM (Data-driven Threat Modelling), presents a comprehensive and innovative approach to data-driven threat modelling, specifically designed to enhance the understanding and differentiation of data-related threats, surpassing existing models in terms of value. d-TM offers distinct advantages stemming from its integration of data across multiple levels of abstraction and phases. By incorporating this comprehensive approach within the organizational architecture, d-TM enables a methodical examination of the attack landscape, extending from the user endpoint to the target data storage. The focus is mainly on the actors involved and the different levels of threat layers that have been identified. d-TM adopts a three-tiered strategy that subdivides data cybersecurity assurance into management, control, and business factors. Each of these factors is viewed as a composite of three interconnected components: storage, processing, and transmission. Moreover, the model leverages its visual presentation of digital assets interaction among each other's using data-flow diagram (DFD) and dynamic capabilities to adapt to the evolving threat landscape, offering the latest updates by interfacing with prominent security catalogues such as MITRE CWE, CAPEC and NIST.

The innovative data-driven threat modelling (d-TM) approach focuses on weaknesses as the root cause of vulnerabilities, which empowers organizations to proactively strengthen security measures. The d-TM model is further empowered by its automation capabilities, which automate the entire process of threat analysis, streamlining and expediting the identification, assessment, and ascertaining the most effective controls to mitigate threats. Finally, d-TM is evaluated using three real case scenarios to determine its applicability to the current emerging industry. The results show that d-TM effectively identifies and quantifies the threats that are potential for any major disruption to the business. The evaluation includes scenarios from healthcare, supply chain and IT service provider sectors. With its emphasis on security assurance and the ability to proactively

address data-related threats, d-TM stands as a practical approach for data-driven threat analysis in ensuring robust cybersecurity.

**Keywords:** *Threat modelling; Data levels; Threat; Weakness; Control; Security assurance; Automation*

# Table of Content

Dedication.....	2
ABSTRACT.....	3
PUBLICATIONS BY THE AUTHOR.....	9
CHAPTER <i>ONE</i> : INTRODUCTION.....	10
1.1 Background.....	11
1.2 Research Challenges.....	12
1.3 Significance of Data-driven Threat Analysis Approach.....	13
1.4 Research Aim and Objectives.....	15
1.5 Research Questions.....	16
1.6 Research Contribution.....	17
1.7 Outline of the Thesis.....	18
CHAPTER <i>TWO</i> : LITERATURE REVIEW.....	20
2.1 Introduction.....	21
2.2 Glossary of Terms.....	21
2.3 Data in Organizations.....	22
2.4 Risk in Data Security.....	25
2.5 Threats Modelling and Analysis Approaches.....	26
2.6 Cyber Threat Intelligence.....	32
2.7 Cybersecurity and Control Standards.....	34
2.8 Data-Driven Threat Modelling Standards and Catalogues.....	36
2.9 Review and Discussion of Related Work.....	36
CHAPTER <i>THREE</i> : RESEARCH METHODOLOGY.....	40
3.1 Introduction.....	41
3.2 Research Methodology.....	41
3.3 Summary.....	48
CHAPTER <i>FOUR</i> : DATA-DRIVEN THREAT MODELLING (d-TM).....	49
4.1 Introduction.....	50
4.2 d-TM Requirements.....	50
4.3 d-TM Fundamental Pillars.....	52
4.4 Conceptual View of d-TM.....	61
4.5 d-TM Process.....	64
4.6 Conclusion.....	83
CHAPTER <i>FIVE</i> : d-TM AUTOMATION.....	85

5.1 Introduction.....	86
5.2 d-TM Platform Overview.....	86
5.3 d-TM Platform Design.....	86
5.4 d-TM Platform Applications.....	88
5.5 Conclusion .....	101
CHAPTER SIX: THE d-TM EVALUATION .....	102
6.1 Introduction.....	103
6.2 d-TM Application: A Real Industrial Case Study (Food Supply Chain).....	104
6.3 d-TM Application: A Real Industrial Case Study (Service-provider).....	114
6.4 Automation Evaluation of d-TM model: A Real Industrial Case Study (Healthcare).....	124
6.5 Comparison with the other works .....	153
CHAPTER SEVEN: THE CONCLUSION.....	157
7.1 Introduction.....	158
7.2 Fulfilling Research Questions.....	158
7.3 Fulfilling Research Objectives.....	159
7.4 The d-TM observed limitation.....	163
7.5 The d-TM Future work .....	164
7.6 Summary.....	165
Appendix A: Set of d-TM Evaluation Questions.....	166
References.....	169

## List of Figures

Figure 1.1. The Research outline .....	19
Figure 2.1. Examples of attack tree applications (Shevchenko <i>et al.</i> , 2018).....	28
Figure 2.2. The Kill Chain stages .....	30
Figure 2.3. PASTA Threat Modelling stages (Shevchenko <i>et al.</i> , 2018).....	31
Figure 2.4. The core constructs of STIX (Barnum, 2014).....	33
Figure 3.1. Methodology of the study.....	44
Figure 4.1. A high-level overview of d-TM. ....	53
Figure 4.2. Data-levels at digital systems .....	54
Figure 4.3. Data levels and phases.....	55
Figure 4.4. d-TM Data levels, layers and actor representation.....	59
Figure 4.5. d-TM conceptual model .....	63
Figure 4.6. Example showing d-TM model components.....	64

Figure 4.7. An overview diagram of the Data-driven threats analysis approach.....	65
Figure 4.8. Business, services and infrastructure relationship mapping.....	67
Figure 4.9. Business interview levels .....	68
Figure 4.10. An example of a d-TM-enabled data flow diagram. ....	73
Figure 4.11. A visual presentation of d-TM weakness and threat analysis .....	75
Figure 5.1 d-TM platform architecture .....	87
Figure 5.2 d-TM platform login, signup and profile page .....	89
Figure 5.3 d-TM platform business data collection interface .....	90
Figure 5.4 Asset details form .....	91
Figure 5.5 Asset administration form .....	92
Figure 5.6 Asset dependency form .....	93
Figure 5.7 Threat analysis functions.....	94
Figure 5.8 Data analysis interface.....	95
Figure 5.9 Weakness Identification Interface .....	96
Figure 5.10 weakness identification table.....	96
Figure 5.11 Threat identification interface. ....	97
Figure 5.12 Threat criticality interface. ....	98
Figure 5.13 Threat criticality outcome table.....	98
Figure 5.14 Threat mitigation interface. ....	99
Figure 5.15 Control assurance interface .....	100
Figure 6.1. d-TM deployed case scenarios .....	103
Figure 6.2. The scenario SAP solution architecture. ....	105
Figure 6.3. A data flow diagram of a business user accessing the SAP application. ....	108
Figure 6.4. The TSS infrastructure topology. ....	115
Figure 6.6. The infrastructure supporting the case study business. ....	126
Figure 6.7. The output table of identified services supporting the hospital business .....	127
Figure 6.8. The output table of infrastructure asset details supporting the patient e-service. ....	129
Figure 6.9. The output table of asset administration details supporting the patient e-service....	130
Figure 6.10. The output table of asset dep. details supporting the patient e-service. ....	131
Figure 6.11. Sample output table of assets data levels and phases for the patient e-service. ....	133
Figure 6.12. The output DFD of the Patient e-service.....	135
Figure 6.13. Sample output table of identified weakness exist in the patient e-service. ....	137
Figure 6.14. Sample output table of identified threats targeting the Careware system. ....	139
Figure 6.15. Sample output table of threat criticalities targeting Careware system. ....	141

Figure 6.16. The output table of threat controls for the careware system. ....	143
Figure 6.17. The output table of control assurance levels for the careware system. ....	144
Figure 6.18. The d-TM dashboard of the case study threat analysis process. ....	145
Figure 6.19. The d-TM visual reports samples for the case study threat analysis process. ....	147

### List of Tables

Table 2.2. STRIDE Threat Categories (Shevchenko <i>et al.</i> , 2018).....	29
Table 2.3. Overview of d-TM model to existing works .....	39
Table 4.1. Mapping CAPEC domains to d-TM infrastructure categories .....	60
Table 4.2. Summary of d-TM Activities.....	66
Table 4.3. Asset type table.....	69
Table 4.4. Asset administration table.....	69
Table 4.5. Asset dependency table.....	70
Table 4.6. A table represents the data level and phase at any asset.....	71
Table 4.7. A d-TM Data flow diagram utilized symbols.....	72
Table 4.8. Business -as- target matrix.....	78
Table 4.9. Threat complexity matrix.....	79
Table 4.10. A table represents the Matrix of threat priority .....	79
Table 4.11. Controls Completeness levels.....	81
Table 4.12. Controls Effectiveness levels.....	82
Table 4.13. Controls Complexity levels .....	82
Table 6.1. Table of critical business processes, services and assets. ....	106
Table 6.2. data level and phase analysis. ....	108
Table 6.3. Threat profile for Bs0. ....	110
Table 6.4. A table representing threat controls to protect (Bs0) data. ....	111
Table 6.6. Business services and processes table .....	117
Table 6.7. Infrastructure details of MSS service.....	118
Table 6.8. Data-levels and phases analysis. ....	120
Table 6.9. Threat analysis of MSS Bs5.....	122
Table 6.10. Threat mitigation and assurance of MSS Bs5.....	123
Table 6.11. Overview of d-TM tool feedback. ....	152
Table 6.12. Overview of d-TM tool to some existing tools .....	155
Table 7.1 Existing threat modelling to d-TM overview(Alwaheidi, Islam and Papastergiou, 2022). .....	161



## PUBLICATIONS BY THE AUTHOR

Date and Authors	Title	Journal/Conference
(Alwaheidi, Islam and Papastergiou, 2022)	“A Conceptual Model for Data-Driven Threat Analysis for Enhancing Cyber Security”	<p><b>Conference Paper</b></p> <p><i>ICR'22 – Proceedings of the ICR'22 International Conference on Innovations in Computing Research pp 365–374.</i></p> <p><a href="https://doi.org/10.1007/978-3-031-14054-9_34">https://doi.org/10.1007/978-3-031-14054-9_34</a></p>
(Alwaheidi and Islam, 2022)	“Data-Driven Threat Analysis for Ensuring Security in Cloud Enabled Systems”	<p><b>Journal Paper</b></p> <p><i>MDPI – Sensors 22 Special Issue Security and Privacy in Cloud Computing Environment, no. 15: 5726.</i></p> <p><a href="https://doi.org/10.3390/s22155726">https://doi.org/10.3390/s22155726</a></p>

# CHAPTER *ONE*: INTRODUCTION

## 1.1 Background

Data is of utmost importance to businesses in the contemporary business world. The importance of data as a strategic asset cannot be understated as firms depend increasingly on digital technology and data-driven decision-making processes. Industry currently experiencing a rise in both the quantity and kind of cyberattacks, twenty-two billion records were made available to the public due to data breaches in the first half of 2021 (Sobers, 2022). 70% of these security breaches were carried out for financial gain, while less than 5% were conducted for the purpose of espionage (Sobers, 2022). On the other hand, hacking was responsible for 22% of the security breaches that were discovered in same year, while phishing and social engineering were responsible for 40% of the breaches, and malware was responsible for 11% of the breaches. According to research conducted by Forbes, it is anticipated that damages caused by cyberattacks would amount to a total of \$6 trillion by the year 2021 (Powell, 2019). Also, it is anticipated that this sum will exceed the total monetary worth of all trading activities that take place globally involving different forms of illegal narcotics (Powell, 2019). In general, the trends in cyberattacks continue to increase, and it is anticipated that these trends will grow even further in the near future. These figures highlight the diverse array of techniques employed by adversaries to exploit vulnerabilities within organizations' systems. Consequently, it is imperative to recognize the upward trajectory of cyberattacks and acknowledge the need for proactive measures to combat this escalating menace.

The integration of technology and digital services in businesses has led to an increase in cyberattacks and incidents. As organizations adopt smart technologies like artificial intelligence, cloud computing, and big data, the attack surface expands, exposing them to new risks (Splunk, 2021). It is crucial for organizations to understand the diverse nature of data generated by technology, encompassing not only business data but also asset-generated data. By focusing solely on business data, organizations may overlook potential security vulnerabilities. A comprehensive understanding of operational data, including asset-generated data, is essential for effective risk mitigation towards ensuring cyber security assurance within the organizational context. Companies invest in digital services and implementing controls, conducting internal security audits and vulnerability assessments that may not be sufficient to provide proper security assurance (Sabillon *et al.*, 2017). However, the evolving nature of cybercrime and the trust placed in the digital realm pose ongoing challenges that organizations need to overcome to protect their business operations (Teoh and Mahmood, 2017). Data plays a vital role in organizational operations, and it is crucial to understand its diverse types and dependencies. Existing techniques often overlook the importance of holistic data understanding in threat analysis. By incorporating a comprehensive approach to data analysis, organizations can better identify and address threats, ensuring the security and integrity of their data assets (Harris *et al.*, 2019).

Threat modelling offers a methodology to help organizations identify and address cyber threats, enabling the implementation of appropriate controls. The proposed d-TM model recognizes three levels of data abstraction: management, control, and business. Additionally, data goes through three phases: rest, use, and motion. Comprehensive analysis of data at different abstraction levels is necessary for effective threat analysis. *Finally, the present research introduces a novel threat modelling technique, known as the data-driven threat model (d-TM), which offers significant value in the field of cybersecurity. The proposed technique encompasses an in-depth understanding of business processes, services, and infrastructure, with infrastructure being classified into five distinct attack levels, ranging from the user agent to data storage. A key emphasis of the d-TM lies*

in its comprehensive treatment of data and its various types during the threat analysis process. By categorizing data and utilizing a Data Flow Diagram (DFD) diagram, common vulnerabilities and threats can be identified, enabling the implementation of suitable control measures to address associated risks. The d-TM model provides a notable contribution by dividing data abstraction into three levels: control, management, and business, along with three phases: data at rest, data in transit, and data in process. Each data level and phase are associated with specific security risks that need to be effectively managed and addressed. Moreover, the d-TM adopts a systematic approach to comprehending organizational data, services, and data flow for thorough threat analysis and control. The prioritization of threats is based on various characteristics, including business-as-a-target scenarios, the complexity of the threat, and its business impact. To enhance the effectiveness of the d-TM, it leverages widely recognized security information from established standards such as Common Attack Pattern Enumeration and Classification (CAPEC), Common Weakness Enumeration (CWE), and NIST SP800-53.

## 1.2 Research Challenges

Ensuring cybersecurity assurance is imperative for organizations to maintain informed decisions for overall business continuity. However, this research has identified several significant challenges pertaining to data threat modelling that hinder the achievement of robust cybersecurity assurance.

- **CHALLENGE 1** *Lack of understanding of various data in the organization*

As highlighted earlier, data stands as the cornerstone of business success. However, data's technical perspective reveals its existence across various levels of abstraction within an organization, often leading to the oversight of crucial data dimensions. Consequently, threat actors exploit vulnerabilities concealed within this data hierarchy. This blind spot poses a remarkable challenge for organizations in effectively fortifying their data assets. Overlooking vulnerabilities across all data abstraction levels augments the risk of potential breaches and compromises. Hence, it becomes imperative for organizations to augment their comprehension of the holistic data landscape, prioritizing the safeguarding of all data components, irrespective of their abstraction level. This approach enables organizations to fortify their overall security stance and ensure the cybersecurity integrity of their data assets. Moreover, from a business standpoint, each organization possesses a diverse array of data types, spanning sales, transactions, income, consumer insights, operations, employee records, financial data, social media interactions, and more. These distinct data forms act as transformative elements for organizations, bestowing invaluable insights for decision-making, strategic plotting, and bolstering operational performance (Alexander, 2017).

- **CHALLENGE 2** *Difficulties of ensuring cybersecurity assurance*

Assurance in the realm of cybersecurity refers to the implementation of security controls to reduce a multitude of potential threats. This assurance stands as a cornerstone for safeguarding organizational assets and upholding the seamless continuity of business operations. Nonetheless, the complex context of organizations renders the landscape replete with dispersed datasets across various organizational segments, causing a formidable challenge. The dispersal of data accentuates the intricacy of establishing comprehensive security controls that envelop the entirety of the data surroundings. To undertake this task effectively, organizations need a coherent grasp of their data's nature, distribution, and associated security vulnerabilities. These intricacies indicate a holistic perspective that recognizes the intricate interconnections among

different facets of organizational data, synchronizes security controls, and acknowledges the essential value of diverse data forms and their potential threats. By assimilating these considerations, organizations can chart a course towards instituting cybersecurity assurance measures that effectively shield their vital data assets and ensure the cybersecurity resilience of their operational foundation. However, faltering to achieve robust cybersecurity assurance within an organization could reverberate as a potential impediment to the overall security investments of its stakeholders(Muscio and Wilson, 2017).

- **CHALLENGE 3** *Limited availability of Effective and automated threat modelling approaches*

Threat modelling is widely used in the industry for security analysis. However, despite a variety of threat modelling approaches, there is still a lack of connection between the analysis of threats based on data and system weaknesses. Understanding the data and weaknesses with data assets to analyse the threats is currently the business's most important need. Thus, a deep comprehension of data intricacies and potential weaknesses becomes paramount for developing robust cybersecurity strategies and safeguarding business continuity. In addition, threat modelling and analysis is a time-consuming and labour-intensive task; therefore, automating threat modelling could increase both effectiveness and adaptability. However, existing threat models have limited automation features available. A lack of emphasis on data and common weaknesses hinders the creation of a realistic threat model.

Organizations typically identify threats through the analysis of cyber breach incidents and respond by implementing the necessary security controls to mitigate the vulnerabilities that were exploited in these incidents(Wood, 2019). However, relying solely on a reactive approach to cybersecurity leaves organizations constantly playing catch-up. To stay ahead of cyber threats, it is imperative to adopt proactive techniques that enable the early identification and prevention of potential threats. This shift in mindset and approach towards proactive cybersecurity measures represents a transformative step in combating cybercrime (Husák *et al.*, 2019).

### **1.3 Significance of Data-driven Threat Analysis Approach**

A thorough analysis of threats is of paramount importance and an urgent necessity for businesses across all sectors. In today's landscape, organizations recognize effective security measures as a critical aspect of ensuring business continuity. With the exponential growth of data being held and generated by businesses from various systems and applications, such as business and operational data, the attack surface expands, and systems become more complex. Consequently, it becomes imperative to analyse this data and undertake necessary measures to prevent it from being compromised by cyberattacks. Therefore, cybersecurity assurance is an immediate demand in any organization to safeguard businesses against such threats.

To underscore the critical need for a data-centric methodology in threat analysis, it is imperative to consider the dynamic landscape of cyber threats that increasingly exploit data as their principal target. The evolving nature of these threats necessitates a nuanced and granular approach to cybersecurity, beyond the scope of traditional asset-based, goal-oriented, or risk-centric strategies. These conventional frameworks, while valuable, may fall short in addressing the intricate and data-

driven environments we navigate today. A data-centric perspective fosters a deeper, more detailed understanding of the data lifecycle, from creation to disposal, thereby fortifying defences against the sophisticated, data-focused threats we now face. This approach is not just a strategic shift but a necessary evolution, aligning with the growing consensus that data is an indispensable asset in the realm of cybersecurity. The imperative for a data-centric approach in cybersecurity is underscored by recent scholarly contributions that highlight the evolving nature of cyber threats and the critical role of data as a target. Turner, McCombie, and Uhlmann (2019) illustrate the effectiveness of a target-centric method in analysing ransomware attacks, emphasizing the significance of focusing on data flows and transactions in cyber intelligence (Turner, McCombie and Uhlmann, 2019). Similarly, Zou et al. (2020) delve into Advanced Persistent Threat (APT) tactics recognition, advocating for a shift towards data-centric security frameworks to bolster Defence mechanisms against sophisticated cyber-attacks (Zou *et al.*, 2020). Taylor, Araujo, and Shu (2020) present a scalable system telemetry model, SysFlow, which encapsulates system activities in a data-centric manner, facilitating in-depth analysis of attack vectors and enhancing cyber threat discovery and forensic capabilities (Taylor, Araujo and Shu, 2020). These state-of-the-art studies not only reinforce the necessity of pivoting towards data-centric security models but also demonstrate their efficacy in addressing contemporary cybersecurity challenges, thereby providing a solid foundation for businesses to adapt and fortify their defences in a data-driven world.

Furthermore, Adopting a data-centric approach to cyber threat analysis presents substantial benefits for businesses, streamlining their cybersecurity efforts and aligning with modern threat landscapes. This methodology enhances threat detection and response capabilities by focusing on protecting data as a critical asset, leading to more efficient resource utilization and time savings. Businesses traditionally employing asset-based or risk-centric models may find the data-centric approach particularly beneficial as it offers a more nuanced understanding of data vulnerabilities and threats, ensuring focused protection efforts. Furthermore, this approach aids in compliance with data protection regulations, thereby reducing legal risks and building customer trust. The strategic focus on data not only improves threat intelligence and incident response but also optimizes effort and resource allocation, allowing businesses to tailor their cybersecurity strategies effectively. By prioritizing the protection of critical data assets, companies can achieve a more targeted Defence mechanism, enhancing overall security posture while supporting business objectives and operational resilience.

Numerous techniques are utilized to ensure the cybersecurity of organizations. However, data-focused threat modelling for cybersecurity assurance has received relatively less attention from researchers, and some existing research may be outdated in addressing emerging technologies and threats associated with the digital shift in today's technology landscape. Moreover, current threat modelling models, such as STRIDE and PASTA, primarily focus on analysing threats based on assets, threats, or actors rather than the specific data being utilized. Additionally, diverse types of data employed by organizations, such as business and operational data, are not adequately covered and distinguished within these models. Furthermore, certain models prioritize vulnerabilities to identify threats and subsequently respond reactively to them. Conversely, the d-TM model concentrates on identifying weaknesses that serve as the underlying cause of the vulnerability. This approach empowers organizations with proactive capabilities to identify potential threats. Despite the release of the National Institute of Standards and Technology's (NIST) Special Publication (SP800-154) on Data-Centric threat modelling in 2016, which focuses on evaluating one type of data (business) for a specific system, it fails to address other types of data, such as operational

data. Moreover, the aforementioned models lack automation, which renders them time-consuming and reliant on human labour.

Considering these gaps, the proposed model d-TM significantly contributes to the threat analysis and security assurance domain and offers dynamic adaptability to modern technologies and granular abstraction levels for organizational data, distinguishing between control, management, and business data. This feature makes it superior to other existing models. Furthermore, the model's data-centric focus provides detailed guidelines for conducting threat analysis based on defined attack surfaces and actors, which sets it apart from other scholarly works in the field. The d-TM model offers a distinct advantage by focusing on weaknesses rather than vulnerabilities used by other models in the threat analysis process. By identifying and analysing weaknesses, the model reveals the root causes that give rise to vulnerabilities, providing a deeper understanding of the underlying issues. Additionally, the d-TM model, as proposed, presents a novel automation capability that facilitates effective and efficient threat analysis procedures. The utilization of automated tools and techniques by the d-TM facilitates the efficient analysis of substantial amounts of data by organizations. This enables the identification of potential threats and the timely implementation of suitable control measures, empowering organizations to ensure their cybersecurity assurance towards the constantly evolving cyber threats.

## 1.4 Research Aim and Objectives

The primary objective of this research is

*to develop an innovative and automated data-driven threat modelling approach that empowers organizations to effectively address data-related cybersecurity threats for overall security assurance.*

To achieve this overarching goal, the research will focus on the following specific objectives:

- **Objective 1.** *Enhance the understanding of organizational data* across various components of the infrastructure, including underlying assets, threats, and weaknesses. This objective aims to provide a comprehensive view of the organization's ecosystem, enabling a holistic assessment of potential weaknesses and threats.
- **Objective 2.** *Develop a data-centric threat analysis modelling approach* that leverages data to evaluate and ensure cybersecurity assurance. This objective entails designing a robust model that incorporates data as a central element in the threat analysis process, providing organizations with valuable insights into threats targeting their data, including a detailed analysis of the underlying weaknesses that contribute to these threats. As well as the most effective controls for mitigation, which assist organizations in strengthening their cybersecurity assurance.
- **Objective 3.** *Validate the effectiveness of the proposed threat modelling approach* through multiple real-world case scenarios. This objective involves applying the developed approach to practical situations, evaluating its performance, and verifying its ability to identify and mitigate threats in diverse operational contexts.
- **Objective 4.** *Develop an automated platform* that streamlines and automates the threat analysis process. This objective focuses on creating a software tool that integrates the data-

driven threat modelling approach, enabling organizations to analyse threats efficiently and effectively, generate actionable insights and the implementation of necessary controls.

By achieving these objectives, this research seeks to contribute to the field of cybersecurity by providing organizations with an advanced and automated approach to threat analysis, empowering them to proactively address data-centric cybersecurity threats and safeguard their critical assets and operations.

## 1.5 Research Questions

This thesis seeks to provide significant contributions to the field of threat modelling by investigating various research questions. Specifically, it aims to shed light on the gaps of current threat modelling approaches, the examination of diverse technical data, the identification and evaluation of threats based on organizational data, the prioritization of threats, and the possible integration of automation to improve threat modelling for comprehensive security assurance.

To address the objectives of this thesis, the following research questions have been formulated:

**RQ1:** *What are the existing gaps in the state of the art with regard to threat modelling approaches?*

This question aims to assess the current state of threat modelling practices and identify any limitations or deficiencies that may exist. By examining the gaps in the current approaches, this study seeks to contribute to the advancement of more comprehensive and effective threat modelling techniques.

*In Relation With: Objective Obj 2, Obj 3*

**RQ2:** *What are the various types of organizational data that need to be analysed?*

This question focuses on understanding the diverse range of organizational data that organizations should consider for analysis. By identifying and categorizing diverse types of data, such as business data, operational data, network logs, and system configurations, this research aims to provide insights into the crucial data sources that play a role in threat identification and analysis.

*In Relation With: Objective Obj 1*

**RQ3:** *How can threats be identified and analysed based on the different data within an organization?*

This question investigates the methodologies and approaches for identifying and analysing threats based on the diverse types of data present in an organization. It aims to uncover potential threats and understand their implications, thereby enhancing the overall threat analysis process.

*In Relation With: Objective Obj 2, Obj 3*



**RQ4:** *How can the identified threats be prioritized to determine the appropriate level of assurance for effective mitigation?*

This question addresses the challenge of prioritizing identified threats. By developing a methodology for threat prioritization, this study aims to enable organizations to allocate resources and efforts more effectively to mitigate the most critical threats and ensure optimal security assurance.

*In Relation With: Objective Obj 2, Obj 3*

**RQ5:** *To what extent can threat modelling be automated to enhance its effectiveness and facilitate wider adoption for overall security assurance?*

This question explores the potential for automating the threat modelling process to improve its efficiency and encourage broader adoption across organizations. By investigating automation techniques and tools, this study aims to assess the feasibility and benefits of automating threat modelling to enhance overall security assurance.

*In Relation With: Objective Obj 3, Obj 4*

## 1.6 Research Contribution

This research makes significant contributions to the field of data-driven threat modelling and aims to enhance cybersecurity assurance for organizations. The proposed model addresses the limitations of existing models and offers a comprehensive approach to mitigate cyber risks associated with data. The unique contributions of this research can be summarized as follows:

**Firstly**, the research introduces a conceptual model that incorporates key concepts related to data in threat modelling, including actors, weaknesses, threats, data, infrastructure, and controls. This conceptual model serves as a foundation for the proposed approach.

**Secondly**, the approach considers data from three distinct levels of abstraction: management, business, and control. It examines these data levels across three phases: at rest, in process, and in transit. This approach provides granular visibility into data at various stages of its lifecycle and explores how attacks can escalate through defined sequential attack layers and actors associated with each data level.

**Thirdly**, the d-TM model offers a comprehensive analysis of data, spanning from actors to data storage. It defines actors and attack surfaces based on the current organizational topology. Furthermore, the model focuses on weaknesses in data assets rather than vulnerabilities, recognizing that weaknesses are the root causes of vulnerabilities. This proactive perspective provides valuable insights into potential threats.

**Fourthly**, the model aims to establish an effective cybersecurity assurance approach by integrating a comprehensive analysis process that encompasses data, weaknesses, threats, controls, and control assurance levels. Additionally, the model incorporates a unique threat prioritization methodology that considers multiple factors to guide the allocation of resources and response strategies.

**Finally**, recognizing the significance of automation in the industry and its role in shaping the future of cybersecurity, this research provides a platform that empowers cybersecurity threat analysis decision-makers to automate the d-TM cyber threat modelling process. By leveraging the power of automation, organizations can enhance their efficiency and effectiveness in identifying and responding to cybersecurity threats.

Overall, this research contributes a novel data-driven threat modelling approach that addresses the limitations of existing models, provides comprehensive analysis and prioritization methodologies, and supports the automation of cybersecurity threat analysis processes. These contributions aim to strengthen organizations' cybersecurity assurance and ensure better protection against evolving cyber threats.

## 1.7 Outline of the Thesis

This section provides an overview of the chapters that comprise the thesis, outlining the flow of the research. The thesis consists of seven chapters, structured as follows:

**CHAPTER 1:** The first chapter discusses the background and motivation for this study and the critical research questions.

**CHAPTER 2:** In this chapter, a comprehensive review of the relevant literature is presented, focusing on threat analysis practices, existing models, and identifying gaps in the literature that motivate the development of the d-TM threat analysis conceptual model.

**CHAPTER 3:** The research methodology employed to address the research objectives and validate the applicability of the proposed model is outlined in this chapter. It discusses the chosen approach, data collection methods, and analysis techniques.

**CHAPTER 4:** The main contribution of this research is presented in this chapter. It presents the *Proposed Data-Driven Threat Modelling (d-TM)*, which includes a conceptual model and outlines the proposed approach, as well as its systemic activities.

**CHAPTER 5:** In this chapter, the architectural layout, functionality, and interface outlooks of the d-TM platform are described and outlined. The chapter provides insights into the technical aspects and design considerations of the platform.

**CHAPTER 6:** This chapter focuses on the evaluation of the d-TM Model. It presents real-world use case scenarios to validate the research, including the methods and procedures employed to gather the necessary information for assessing the model's ability to fulfil the research aims and objectives.

**CHAPTER 7:** This Chapter concludes the research and presents features, limitations, and future work.

These chapters collectively form the structure of the thesis, presenting a comprehensive exploration of the research topic, methodology, proposed model, evaluation, and platform architecture. However, the following figure represents the outline of the d-TM research.

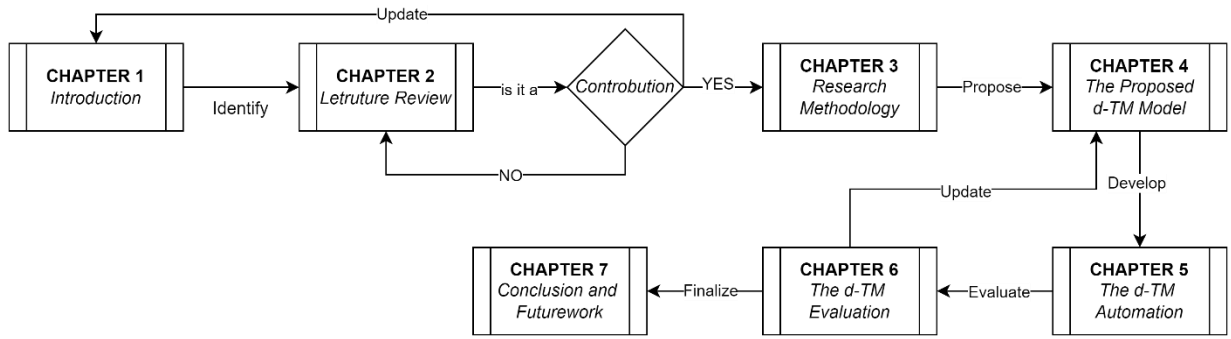


Figure 1.1. The Research outline

## CHAPTER *TWO*: LITERATURE REVIEW

## 2.1 Introduction

This chapter aims to establish a comprehensive understanding of the essential aspects within the research area of data-driven threat analysis modelling for ensuring cybersecurity assurance. By carefully considering key terms such as threat models, cyber security assurance, IT data, threats, and standards for conducting the literature review, we have attempted to provide a solid foundation for the subsequent sections of this thesis.

To begin, the chapter will drill into the concept of threat models, exposing their significance in the field of cybersecurity assurance. A clear understanding of threat models is crucial for comprehending the subsequent discussions on data-driven approaches and their application in threat analysis modelling. Furthermore, this chapter will explore the overarching concept of cyber security assurance, emphasizing its critical role in safeguarding information systems and data from potential threats. By investigating established frameworks and best practices, we will establish a context for the utilization of data-driven approaches within the realm of cybersecurity assurance. Additionally, this chapter will highlight the importance of IT data as a fundamental resource in threat analysis modelling. Understanding the nature and characteristics of these data sources is pivotal in effectively applying data-driven techniques for threat analysis.

Moreover, the chapter will provide a thorough review of relevant literature to identify similar works and approaches in the field. By examining existing research, and studies that align with the approach pursued in this work, we will establish a foundation for building upon and contributing to the existing body of knowledge. This review will not only demonstrate the significance of the proposed research but also highlight potential gaps and areas for further exploration. Through this comprehensive exploration of the research area, this chapter aims to provide a solid framework and collective understanding of the essential concepts and terminology related to data-driven threat analysis modelling for ensuring cybersecurity assurance. By contextualizing the subsequent sections of this thesis within the existing body of knowledge, we aim to establish a robust foundation for the research presented in this work.

## 2.2 Glossary of Terms

To maintain clarity in this thesis, it is important to establish a common understanding of specialized terminology in threat modelling. In this subsection, we aim to demystify specific terminology and key concepts for our readers. This will provide a coherent context for engaging with our research and improve comprehension. Table 2.1 serves as a dictionary for the thesis, outlining fundamental terms for threat analysis. The glossary will help readers navigate the technical landscape explored in this thesis.

Term	Definition
IT	Information Technology
IoT	Internet of Things
MITM	Man in the middle attack
VM	Virtual Machine
RDP	Remote Desktop Protocol
BGP	Boarder Gateway Protocol
FTP	File Transfer Protocol

TTP	Tactics, Techniques, and Procedures
DFD	Data-flow Diagram
SSH/TELNET	Secure Shell/teletype network
Netconf	Network Configuration Protocol
LAN	Local Area Network
WAN	Wide Area Network
IoC	indicators of compromise
D/DoS	distributed denial of service
API	application programming interface
	Domain Name Service

Table 2.1. Glossary of terms

## 2.3 Data in Organizations

Organizational data plays a crucial role in shaping the characteristics of most organizations. It improves decision-making by using methods and data-driven approaches. By extracting insights and identifying patterns from the data, organizations can make informed and accurate decisions (Sharma, Poulouse and Maheshkar, 2023). However, Businesses encounter a range of obstacles when utilizing data. One prominent hurdle arises from concerns surrounding security and privacy (Aloufi and Abdulaziz, 2022). Additionally, businesses may face difficulties in the technological realm, particularly in terms of their limited knowledge of data and challenges in adopting novel analytical tools (Zulkarnain *et al.*, 2021). In a broader sense, harnessing organizational data through analytics and data-driven approaches can yield enhanced decision-making capabilities, heightened efficiency, and more favourable business outcomes. It is obvious that data type ranges are increasing with the use of modern approaches to data visualization, description and modelling (McNulty, 2022). However, organizational data can be further categorized into subcategories, including business data, Information Technology (IT) data, Big data, Internet of Things (IoT) data, and blockchain data. Regardless of the specific context in which a particular class of data is utilized, ensuring data security is a fundamental requirement.

Understanding data in organization could be difficult, due to the variety of incorporated data to run businesses. Hence, the research abstracted data in cyber realm into two distinct types: Technology-related and Business-related data. The distinction between these two types of data is crucial for organizations to effectively manage their information assets and support both operational efficiency and strategic decision-making. Technology-related data is centered around information pertinent to infrastructure assets, encompassing elements like communication protocols and administrative access points. This type of data is crucial for managing and safeguarding the technological backbone of an organization, ensuring that systems and networks function seamlessly and securely. On the other hand, business data delves into the core operational elements of an organization, dealing with critical information such as financial records, customer interactions, and employee details. This data is fundamental to strategic decision-making and day-to-day operations, driving growth and efficiency within the organization. These broad categories of data abstraction can be further refined and segmented to achieve greater specificity and relevance in particular contexts. By subdividing technology-related and business data into more granular categories, organizations can tailor their data management and analysis practices to meet specific needs, enhance precision in decision-making, and address unique challenges inherent to

their operational environment. However, organization should adopt a comprehensive cybersecurity framework that consider data and its varieties to ensure their cybersecurity posture.

- **Technology-Related Data (TRD)**

The digitalization of business operations and the reliance on digital infrastructure make organizations vulnerable to cybersecurity threats, including malware, ransomware, phishing attacks, and data breaches. Kostayeva and Chemyakov (2020) discuss how the advancement of digital technologies has escalated cyber-related risks, emphasizing the importance for businesses to take proactive measures against these risks to protect their data and ensure business continuity (Ntsiepdjap, 2022). The management of technology-related such as generated, processed, and stored data presents significant challenges from a cybersecurity perspective, encompassing various aspects such as data integrity, confidentiality, cyber-physical system security, and trust. Ensuring the integrity and confidentiality of vast quantities of data generated by technological systems is a paramount challenge, particularly in sensitive domains like mental healthcare, where the risks associated with cybersecurity breaches are amplified due to the sensitive nature of the data and the potential vulnerability of data donors, as discussed by (Ive, 2022). Additionally, Cyber-Physical Systems (CPS), such as interconnected autonomous vehicles, pose unique cybersecurity challenges due to their reliance on real-time data for safe operation. (Schmittner *et al.*, 2019) highlight the critical importance of ensuring the availability, integrity, authenticity, and accountability of data in CPS to prevent life-threatening situations arising from successful cyber-attacks. Furthermore, the proliferation of IoT devices and their integration into critical systems necessitates a focus on trustworthiness. (Lee *et al.*, 2022) identify critical challenges in IoT value creation related to cybersecurity, including continuous scaling-up of systems, co-creation, data-driven value creation, and user-centric design, which are intertwined with trust factors central to the value creation process in IoT platforms and systems. On other side, these assets are operated and design by human, that could imply another security risk. For example, The impact of asset takeover due to misconfiguration or improper configuration or design is a critical issue in cybersecurity, exposing organizations to significant risks. Misconfigurations can lead to vulnerabilities in web applications, undermining various sectors including finance, healthcare, and Defence, as Alkahla, Shatnawi, and Taqieddin (2021) emphasize in their study on web security vulnerabilities. They point out that misconfigurations result from flaws in design, implementation, operation, or management, affecting different levels of a web application (Al-Kahla, Shatnawi and Taqieddin, 2021). These studies underscore the complexity of managing configurations in the digital infrastructure and the need for advanced tools and methodologies to identify, report, and mitigate misconfigurations effectively, thereby reducing the risk of asset takeovers and enhancing the overall security posture of organizations. However, The research in the proceeding sections will discuss more granular abstraction to technology-related data to provide more insights and efficacy to mitigate particular threats, such as account takeover discussed by (Al-Kahla, Shatnawi and Taqieddin, 2021). In summary, Addressing Technology-related data challenges requires a comprehensive approach that includes the development of robust data protection measures, and the fostering of trust through transparent and secure data practices, necessitating ongoing vigilance and innovation in cybersecurity strategies as technology continues to evolve.

- **Business-Related Data (BRD)**

Business data represents a fundamental resource that shapes an organization's strategies and operations. It encompasses information retrieved and stored by an organization for various

purposes. Due to the valuable insights, it contains, business data requires elevated levels of security to prevent unauthorized access by malicious individuals or systems. Running business data on an organization's digital infrastructure exposes it to various risks that can have significant consequences for the business's operations, reputation, and financial standing. These risks are multifaceted and stem from both internal and external sources. Business related data could be any data that participate in organization day to day business operation, such financial, customer, or employees. Financial data forms the backbone of business data analysis and is primarily useful to major stakeholders such as administrators, regulators, and investors. financial data is defined as information that describes a company's financial background and its performance as measured by financial metrics(Farboodi *et al.*, 2022). This information is crucial for assessing the effectiveness of an organization's strategies in meeting performance and developmental goals. Financial data enables decisions regarding the adoption or replacement of strategies based on their demonstrated effectiveness. In addition to internal stakeholders, external parties use reported financial data to evaluate credit ratings and make investment decisions. Financial data encompasses various categories depending on the specific aspects of the business(Amit and Schoemaker, 1993). Examples include assets (real property, private property, and intangible property), liabilities (financial obligations to lenders and creditors), and equity (residual possessions and belongings after repaying debts using available assets). Equity comprises shares, common stock, and preferred stock. Overall, a company's financial information may contain insights into its operational strategies. While customer data is another category of organizational information that requires protection to prevent the disclosure of a company's competitive strategies. According to Deshpande (2020), customer data refers to the information customers provide when interacting with business organizations through various channels such as websites, mobile applications, surveys, social media, and marketing campaigns, both online and offline. This data provides valuable insights into customers' behavioural, personal, and demographic characteristics (Deshpande, 2020). Customer information can be further categorized into personal data, engagement data, attitudinal data, and behavioural data, allowing organizations to better understand and serve their customer base. Also, Employee data is equally important as customer data, as both have a direct impact on organizational performance. According to Jodka (2018), employee data comprises information such as payroll details, leave records, medical information, and other relevant data related to the workforce. Employee data encompasses all the information that employers collect about their employees and serves as a crucial component of an organization's performance strategy(Jodka, 2018). Protecting employee and customer data is of paramount importance to organizations. Safeguarding this data from unauthorized access, breaches, or misuse is crucial to maintaining trust, privacy, and legal compliance. Organizations need robust security measures, including access controls, encryption, and monitoring systems, to ensure the confidentiality, integrity, and availability of employee and customer data.

In conclusion, Cybersecurity incidents are ranked as the number one risk threatening business continuity, with costly data breaches being significant challenges for any organization. The technology, while beneficial, also expose businesses to threats such as cyber-attacks that exploit vulnerabilities in applications and systems (Hytönen, Trent and Ruoslahti, 2022). Due to that, organization should have comprehensive understanding of running data to have informed decisions about security risks. Data in an organization, comprising diverse categories of data that possesses substantial significance for the continuous operation of a business. Data pertaining to technology and data pertaining to business are the fundamental components of these investigations, with technology data being further categorized into two distinct types or levels (as employed within the



study framework): control and management, in conjunction with business data serving as a third level of data. However, Effective management and security of these data are imperative to maintain operational efficiency, support decision-making processes, and safeguard the competitive position of the organization.

## 2.4 Risk in Data Security

In the cybersecurity realm, risk analysis and threat analysis are two critical components that, although related, serve different purposes and involve distinct processes. Risk analysis in cybersecurity is a comprehensive process that involves identifying, assessing, and prioritizing risks to an organization's information assets. This process takes into account the potential impact of identified risks on the organization's operations and objectives. The goal of risk analysis is to inform decision-making regarding which cybersecurity measures should be implemented to mitigate identified risks effectively. It considers various factors, including the likelihood of risk occurrence, the vulnerability of assets, and the potential impact of risk realization on the organization. Risk analysis is pivotal in developing a risk management strategy that balances the cost of protective measures with the benefits of risk reduction (Insua *et al.*, 2019; Portalatin *et al.*, 2021). While threat analysis, on the other hand, focuses specifically on identifying, categorizing, and assessing potential threats to an organization's cybersecurity. This includes analysing the capabilities, intentions, and methods of potential threat actors, as well as identifying the vulnerabilities and weaknesses they might exploit. Threat analysis aims to understand the threat landscape the organization faces and to prepare for or mitigate potential attacks. By understanding the nature and source of potential threats, organizations can tailor their cybersecurity defences more effectively to protect against specific types of attacks (Kaja, Shaout and Ma, 2019; Luo *et al.*, 2021). The main difference between the two lies in their scope and focus: risk analysis is broader, encompassing all types of risks (including but not limited to cybersecurity threats) and their impact on the organization as a whole, while threat analysis is more narrowly focused on the cybersecurity domain, concentrating on the threats themselves and their characteristics.

In today's digital landscape, organizations face a wide range of cybersecurity threats that can have severe consequences if not adequately addressed. Kohen (2019) highlights several common forms of cybersecurity threats, including accidental sharing, overworked cybersecurity teams, ransomware, poor password hygiene, phishing emails, fraud, and denial. The increased dependence on computers, networks, and social media has significantly contributed to the rise in cybersecurity risks across the globe. Accidental sharing of sensitive information, either through human error or improper data handling, can lead to unintended exposure of confidential data. Overworked cybersecurity teams may struggle to keep up with the rapidly evolving threat landscape, making organizations more vulnerable to attacks. Ransomware, malicious software that encrypts data and demands a ransom for its release, has become a pervasive threat targeting organizations of all sizes. Poor password hygiene, such as using weak passwords or reusing them across multiple accounts, creates opportunities for cybercriminals to gain unauthorized access to systems and sensitive data. Phishing emails, disguised as legitimate communications, trick users into revealing sensitive information or downloading malicious attachments. Fraudulent activities, including identity theft and financial fraud, pose significant risks to individuals and organizations alike. Denial-of-Service (DoS) attacks, which aim to overwhelm a system or network to disrupt its normal functioning, can cause severe disruptions to businesses. To mitigate these risks, organizations must implement robust security measures, including encryption, multi-factor authentication, employee training on cybersecurity best practices, and regular system patching and

updates. Additionally, establishing incident response plans and conducting regular vulnerability assessments and penetration testing can help identify and address potential weaknesses before they are exploited.

In conclusion, the ever-increasing cybersecurity risks in the digital landscape necessitate proactive measures to protect sensitive data and ensure the integrity of organizational systems. Understanding and managing these risks requires a comprehensive approach that encompasses technological solutions, employee awareness and training, and collaboration between various stakeholders. By adopting effective risk management strategies, organizations can enhance their cybersecurity assurance and minimize the potential impact of cyber threats on their operations and reputation.

## **2.5 Threats Modelling and Analysis Approaches**

In the cyber threats realm, it is imperative to understand the difference between threat modelling and threat analysis. Threat modelling is a proactive, systematic process used during the design phase to identify and assess potential threats and vulnerabilities within a system's architecture(Alsmadi, 2019). In contrast, threat analysis is a broader term that encompasses the identification, evaluation, and prioritization of threats across an organization's entire information ecosystem (Wolf and Serpanos, 2019). Furthermore, threats could be imposed due to vulnerability or a weakness in the system that could be exploited by attackers. understanding the difference between weaknesses and vulnerabilities is crucial. Weaknesses are general flaws or deficiencies that could potentially lead to security issues but have not yet been exploited. Vulnerabilities are specific, identifiable security flaws that can be directly exploited by threat actors(Wolf and Serpanos, 2019). However, the following methodologies and distinctions form the foundation of effective cybersecurity strategies, enabling organizations to address both current and emerging threats comprehensively.

### **2.5.1 Threat Analysis Approaches**

In the realm of cybersecurity, different types of threat analysis methodologies such as risk-based, goal-based, and asset-based analyses are employed to identify, evaluate, and prioritize potential threats. Each type of analysis offers a unique perspective and approach to understanding and mitigating threats.

**2.5.1.1 Risk-Based Analysis** focuses on evaluating threats based on the likelihood of their occurrence and the potential impact on the organization. This approach helps prioritize risks and allocate resources effectively to address the most significant threats. Risk-based cyber threat analysis can be used to improve cybersecurity by providing valuable insights and understanding of the patterns and relationships in cyber threats faced by organizations (Pires and Mascarenhas, 2023). This analysis involves exploring data, performing hypothesis testing, and using correlation techniques to identify noticeable patterns and validate their presence (Liu *et al.*, 2023). By analysing cyber threat intelligence, decision-makers can reduce uncertainty and improve the accuracy of risk analysis, leading to more informed decision-making (Dekker and Alevizos, 2024). This approach considers both known unknowns and unknown unknowns, utilizing methodologies such as causal graphs to reduce uncertainty and improve predictability(Dekker and Alevizos, 2024).

**2.5.1.2 Goal-Based Analysis** centres on the organization's objectives, identifying threats that could hinder achieving these goals. This method aligns security measures with the organization's strategic aims, ensuring that protective efforts support overall business success. This can be achieved through the use of goal recognition algorithms, which analyse attack graphs to identify the objectives of actors in a computer network (Zhang *et al.*, 2022). These algorithms address the challenges of dealing with noisy and partial observations, as well as the need for fast, near-real-time performance (Zhang *et al.*, 2022). By improving the accuracy and runtime of goal recognition, these algorithms can enhance risk management and alert correlation mechanisms for intrusion detection (Jiang *et al.*, 2023).

**2.5.1.3 Asset-Based Analysis** targets specific organizational assets, such as data, hardware, and software, assessing threats directed at these resources. This approach prioritizes the protection of critical assets, ensuring their integrity and availability. Asset-based cyber threat analysis has several advantages. It allows for the identification and evaluation of specific assets that are at risk, enabling organizations to prioritize their security measures and allocate resources effectively (Kawanishi *et al.*, 2023). Additionally, asset-based analysis helps in understanding the impact of cyber threats on the organization's operations and the economic cost of implementing security measures (Shen *et al.*, 2022). It also enables the detection of potential dangers to software systems, allowing developers to add mitigations and enhance the dependability and safety of their designs (K, T and V, 2023). However, asset-based analysis may not capture all possible threats and vulnerabilities, as it focuses primarily on the identified assets and their connections. It is important to complement asset-based analysis with other threat modelling methods to ensure comprehensive coverage of potential risks.

## **2.5.2 Threat Modelling Approaches**

Threat analysis models reveal a significant body of work dedicated to enhancing cybersecurity through systematic threat identification, analysis, and mitigation strategies. The continuous evolution of threat analysis models in cybersecurity reflects the field's dynamic nature and the ongoing efforts to develop more effective, efficient, and context-aware strategies to counteract sophisticated cyber threats. The integration of traditional models with modern technologies and methodologies marks a significant advancement in the cybersecurity domain, offering promising avenues for future research and application in safeguarding digital assets and infrastructures. Attack Tree, STRIDE, PASTA, and Kill Chain, These models serve as foundational frameworks within the cybersecurity domain, guiding the assessment of potential threats and the development of robust Defence mechanisms.

### **2.5.2.1 The Attack Tree**

The attack tree is regarded as one of the oldest approaches to threat modelling. Despite being considered an old model, the attack tree remains one of the most widely employed methods, even in the modern world. It is mostly applied in cyber-only systems, cyber-physical systems, and purely physical systems (Shevchenko *et al.*, 2018). Initially, attack trees were used mainly as standalone methods. However, the technique has since been combined with other models to make it more effective (Shevchenko *et al.*, 2018). Attack trees occur in the form of diagrams that are used to depict potential attacks on a system (Shevchenko *et al.*, 2018). The model is designed in the form of a tree such that the root of the tree shows the goal of the attack, while its leaves are the methods through which the goal is to be realized (Shevchenko *et al.*, 2018). Figure 2.1 illustrates

an example of an attack tree application; this system results in a situation whereby the threat analysis occurs in the form of a tree.

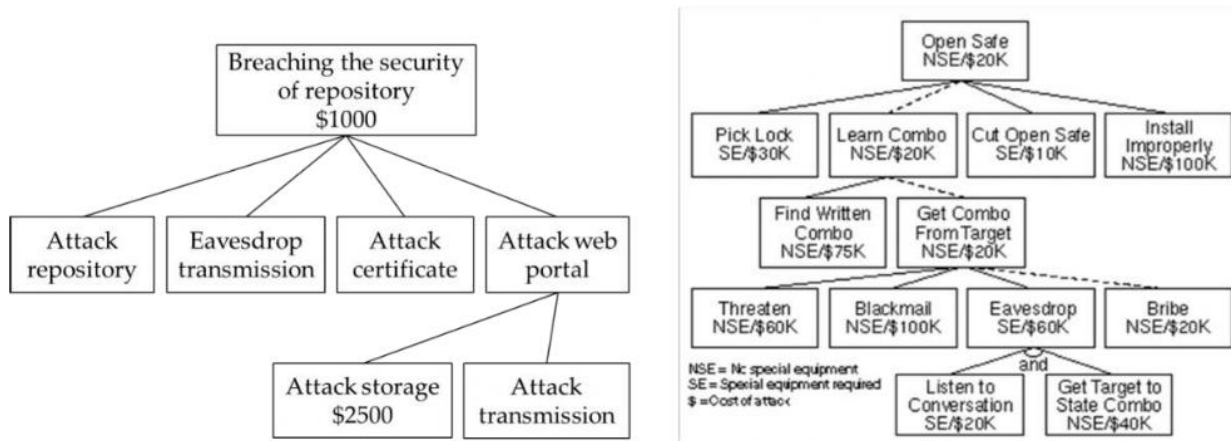


Figure 2.1. Examples of attack tree applications (Shevchenko *et al.*, 2018)

### 2.5.2.2 The STRIDE Model

The STRIDE model, invented in 1999 and later adopted by Microsoft Corporation in 2002, has established itself as one of the most sophisticated methods of threat modelling (Shevchenko *et al.*, 2018). Over time, the STRIDE model has evolved, incorporating various enhancements such as the use of threat-specific tables and additional variants like STRIDE-per-integration and STRIDE-per-Element (Shevchenko *et al.*, 2018). These evolutionary changes have expanded the model's capabilities and made it more adaptable to different contexts and scenarios.

The mechanism of the STRIDE model primarily revolves around the detailed design of a system (Shevchenko *et al.*, 2018). By constructing data-flow diagrams, the STRIDE model enables the identification of entities, events, and system boundaries, facilitating a comprehensive understanding of the system's architecture (Shevchenko *et al.*, 2018). This analysis helps in recognizing potential vulnerabilities and threats that may arise from the system's design and structure.

A key aspect of the STRIDE model is its application of a set of known threats based on mnemonic identities. These mnemonic identities represent diverse types of threats that a system may face. The six mnemonic identities used in the STRIDE model are as follows:

- **Spoofing:** This refers to an attack where an unauthorized entity impersonates another entity to gain access to sensitive information or perform malicious activities.
- **Tampering:** Tampering involves the unauthorized modification or alteration of data, systems, or processes with malicious intent.
- **Repudiation:** Repudiation attacks involve one party denying their actions or transactions, making it difficult to hold them accountable.
- **Information Disclosure:** This type of attack occurs when sensitive information is accessed or disclosed to unauthorized parties.

- Denial of Service: Denial of Service attacks aim to disrupt or disable a system, network, or service, rendering it inaccessible or unusable for legitimate users.
- Elevation of Privilege: Elevation of Privilege attacks involve an unauthorized entity gaining higher privileges or access rights within a system or network, enabling them to perform actions beyond their authorized scope.

By considering these mnemonic identities, the STRIDE model provides a systematic approach to identify and categorize potential threats, enabling organizations to prioritize their mitigation efforts based on the severity and potential impact of each threat. Table 2.2 illustrates the STRIDE threat categories, visually representing these mnemonic identities and their corresponding threats within the model (Shevchenko *et al.*, 2018). This visualization enhances the understanding of the STRIDE model and facilitates its application in threat analysis and mitigation strategies.

	Threat	Property Violated	Threat Definition
S	Spoofing identity	Authentication	Pretending to be something or someone other than yourself
T	Tampering with data	Integrity	Modifying something on disk, network, memory, or elsewhere
R	Repudiation	Non-repudiation	Claiming that you didn't do something or were not responsible; can be honest or false
I	Information disclosure	Confidentiality	Providing information to someone not authorized to access it
D	Denial of service	Availability	Exhausting resources needed to provide service
E	Elevation of privilege	Authorization	Allowing someone to do something they are not authorized to do

Table 2.2. STRIDE Threat Categories (Shevchenko *et al.*, 2018)

In summary, the STRIDE model has evolved into a sophisticated method of threat modelling since its inception. By evaluating system design, utilizing mnemonic identities, and employing data-flow diagrams, the STRIDE model enables organizations to identify, categorize, and prioritize threats systematically. This approach enhances the overall cybersecurity posture by directing efforts toward the most critical vulnerabilities and facilitating the development of effective countermeasures.

### 2.5.2.3 The Kill Chain

Cyber kill chains are another popular method of threat modelling that is used in different sectors involving cyber systems. It is a collection of processes that are related to the use of cyberattacks on various systems (Bala Bharathi and Suresh Babu, 2018). According to Bharathi & Babu (2018), the kill chain describes the stages that constitute a successful cyberattack (Bala Bharathi and Suresh Babu, 2018). This model is a stepwise description of the mechanism of complex attacks (Bala Bharathi and Suresh Babu, 2018). Conventionally, a kill chain reference in an indicator shows that the indicator detects malicious behaviours at a given phase of the kill chain (Bala Bharathi and Suresh Babu, 2018). For instance, a kill chain reference in a TTP shows that such a TTP is used at the considered phase of the kill chain (Bala Bharathi and Suresh Babu, 2018). One of the most common forms of the cyber kill chain model involves the use of seven distinctive steps proposed by the Lockheed Martin Defence company (Bala Bharathi and Suresh Babu, 2018).

The main phases of this model are reconnaissance, weaponization, delivery, exploitation, installation, command and control, and actions on objectives (Bala Bharathi and Suresh Babu, 2018). Each of the steps is characterized by specific data types depending on the aspects of the associated activities. For instance, reconnaissance may be associated with data relating to web analytics or firewall penetration (Bala Bharathi and Suresh Babu, 2018). According to Bharathi and Babu (2018), precaution is one of the most vital aspects of the kill chain model. It involves taking measures before the anticipated attack and initiating quick responses in the event of the detected invasion (Bala Bharathi and Suresh Babu, 2018). The mechanism requires a proper understanding of the orchestrators of the anticipated attacks (Bala Bharathi and Suresh Babu, 2018). The kill chain model describes the specific attack phases of assaults from the perspectives of cybercriminals (Bala Bharathi and Suresh Babu, 2018). This information offers the opportunity to understand vulnerabilities from the attacker’s perspective and facilitates the determination of the actions that are best suited to counteract the expected attacks (Bala Bharathi and Suresh Babu, 2018). Each of the seven stages of threat modelling in the kill chain model is intended to undertake a specific objective.

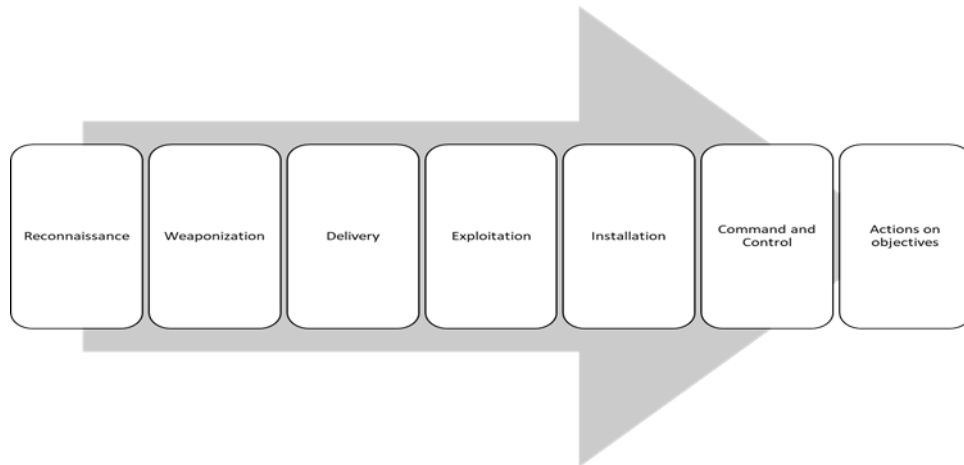


Figure 2.2. The Kill Chain stages

#### 2.5.2.4 The Process for Attack Simulation and Threat Analysis (PASTA)

PASTA (Process for Attack Simulation and Threat Analysis) is a relatively recent advancement in the field of risk modelling. Developed in 2012, PASTA represents a risk-centric model of threat modelling that offers a structured approach to assess and address potential threats (Shevchenko *et al.*, 2018). Like the kill chain model, Figure 2.4 represents the PASTA model, which consists of seven distinct stages, each encompassing multiple activities that serve various purposes in the processes of threat prediction and counteraction.



Figure 2.3. PASTA Threat Modelling stages (Shevchenko *et al.*, 2018)

The division of the PASTA model's functional mechanism into stages makes it highly adaptable and easy to implement in different contexts. The primary objective of the PASTA model is to consolidate business objectives and requirements, ensuring that threat modelling aligns with the strategic goals of the organization (Shevchenko *et al.*, 2018). Each stage of the PASTA model involves the utilization of a wide range of design and elicitation resources, allowing for a comprehensive analysis of potential threats (Shevchenko *et al.*, 2018).

One key aspect of the PASTA model is the inclusion of key decision-makers from various departments and the incorporation of security input from operations, architecture, governance, and development. This multi-disciplinary approach ensures that threat modelling is conducted from a holistic perspective, considering different facets of the organization's structure and operations (Shevchenko *et al.*, 2018). By involving key stakeholders and subject matter experts, the PASTA model elevates the threat modelling process to a strategic level, enhancing its effectiveness and relevance to the organization's overall cybersecurity strategy. Furthermore, the PASTA model adopts an attacker-centric approach to provide an asset-centric output. It analyses anticipated threats from the perspective of potential attackers, enabling organizations to gain valuable insights into their vulnerabilities and prioritize mitigation efforts accordingly (Shevchenko *et al.*, 2018). This approach allows for more targeted threat scoring and enumeration, helping organizations develop high-efficiency threat detection systems that are specifically tailored to their unique assets and risk landscape (Shevchenko *et al.*, 2018). However, PASTA represents an innovative and

comprehensive risk-centric model of threat modelling. By incorporating multiple stages, involving key decision-makers, and adopting an attacker-centric approach, PASTA enhances the accuracy and effectiveness of threat analysis. It empowers organizations to align threat modelling with their strategic objectives, effectively manage risks, and develop robust cybersecurity measures to protect their valuable assets and systems.

## 2.6 Cyber Threat Intelligence

Cyber Threat Intelligence (CTI) plays a crucial role in combating cybercrime by providing organizations with valuable information to diagnose and understand the threats they face. It encompasses the collection of data that helps companies comprehend past, present, and future threats, enabling them to prepare, prevent, and identify risks (Tounsi, 2019). By leveraging cyber intelligence, organizations can establish robust Defence systems to mitigate potential threats that could have detrimental consequences (Tounsi, 2019).

One significant aspect of cyber threat intelligence is the collection of raw data that describes existing or potential threats. This data undergoes analysis to develop actionable solutions and automated security controls (Tounsi, 2019). The primary objective of cyber threat intelligence is to keep organizations well-informed about common types of risks and effective countermeasures to mitigate them. Several categories of cyber threat intelligence exist, including Structured Threat Information Expression (STIX), Open-source intelligence (OSINT), and Geospatial Intelligence (GEOINT).

**2.6.1 The Structured Threat Information Expression (STIX)** is a standardized language developed by the OASIS Cyber Threat Intelligence (CTI) Technical Committee and MITRE to describe cyber threat information (Barnum, 2014; Sadique *et al.*, 2018). It allows for consistent sharing, storage, and analysis of threat information. STIX is intended for various parties involved in system protection, such as cyber defenders, threat analysts, malware analysts, and security researchers (Barnum, 2014). The language provides a universal framework for describing threats, promoting efficient communication, and facilitating the automation of threat detection and prevention activities (Sadique *et al.*, 2018).

The STIX language consists of nine principal constructs that work together to simplify the description of threat information (Barnum, 2014; Sadique *et al.*, 2018). These constructs include observables, incidents, exploit targets, indicators, reports, threat actors, Adversary Tactics, Techniques, and Procedures (TTPs), courses of action, and campaigns (Barnum, 2014; Sadique *et al.*, 2018). Each construct serves a specific purpose in capturing and conveying threat-related information. For example, observables represent objects seen or to be seen in cyber systems, while indicators define patterns and meanings derived from observables (Barnum, 2014; Sadique *et al.*, 2018). Incidents refer to adversary actions, and TTPs encompass the attack patterns employed by cybercriminals (Barnum, 2014; Sadique *et al.*, 2018). Exploit targets denote vulnerabilities at risk of exploitation, while courses of action are responses to these attacks (Barnum, 2014; Sadique *et al.*, 2018). Reports collect relevant STIX content and facilitate shared information. Figure 2.5 illustrates the core constructs of STIX, providing a visual representation of their relationships (Barnum, 2014).



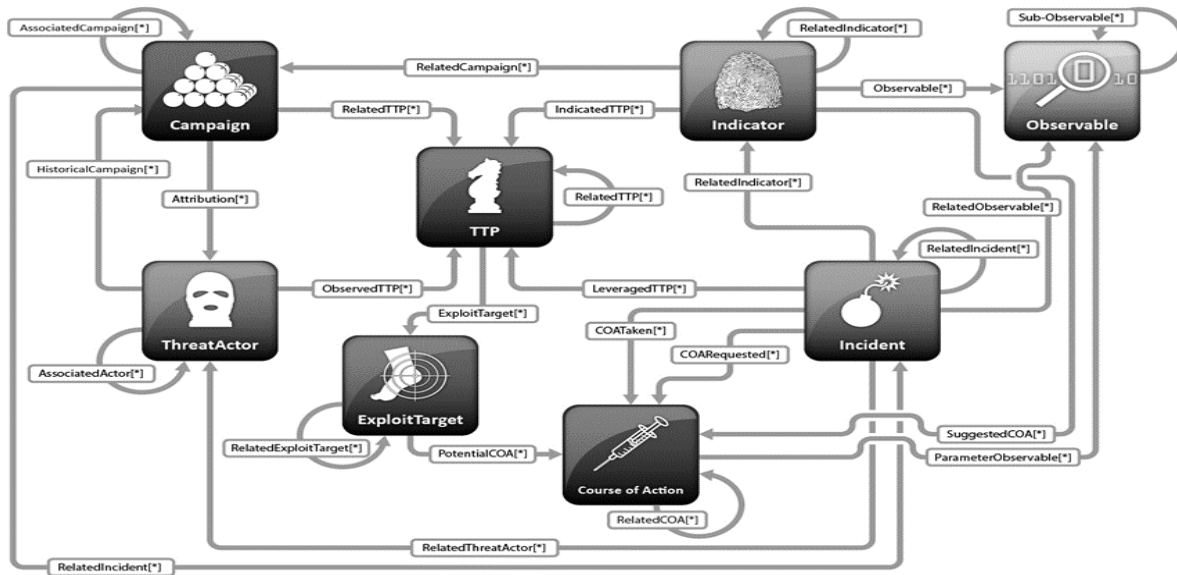


Figure 2.4. The core constructs of STIX (Barnum, 2014).

**2.6.2 Open-source intelligence (OSINT)** is another category of cyber threat intelligence that gathers data from publicly available sources, such as social media reports, news, and public reports (Impe, 2018). Unlike open-source software, OSINT focuses on collecting information from various public sources and has become increasingly valuable in computer security (Impe, 2018). Social media intelligence, a subset of OSINT, plays a significant role in extracting relevant information from social media platforms.

**2.6.3 Geospatial Intelligence (GEOINT)** is an approach to cyber threat intelligence that gathers information from geospatial data sources like maps and GPS systems (Impe, 2018). By incorporating geospatial data, GEOINT provides additional insights into the geographical contexts of cyber threats (Impe, 2018). However, it is important to exercise caution when using GEOINT data for geographical attribution due to potential false flags and uncertainties (Impe, 2018). GEOINT aims to introduce a geospatial dimension to threat detection and prevention activities, enhancing situational awareness.

The applications of Cyber Threat Intelligence are diverse and encompass various aspects of an organization's cybersecurity strategy. CTI plays a vital role in proactive Defence by enabling organizations to identify and prioritize potential threats based on their relevance and severity (Tounsi, 2019). It assists in the development of incident response plans, threat-hunting activities, and vulnerability management processes. Moreover, CTI supports the identification of indicators of compromise (IOCs) that can be used to detect and respond to ongoing attacks (Tounsi, 2019). By monitoring IOCs, organizations can detect malicious activities and take timely actions to mitigate the impact of an attack.

In conclusion, cyber threat intelligence is a vital component in the fight against cybercrime. STIX, as a standardized language, facilitates consistent communication and sharing of threat information. OSINT gathers data from public sources, while GEOINT incorporates geospatial data for enhanced

threat analysis. By leveraging these cyber threat intelligence categories, organizations can strengthen their security measures and proactively defend against evolving cyber threats.

## **2.7 Cybersecurity and Control Standards**

Cybersecurity standards are crucial for organizations of all sizes and categories, providing a set of best practices established by experts to protect against cyber threats (Collier *et al.*, 2014). These standards outline key measures and implementation guidelines to ensure effective protection against cybercrime. Some of the primary cybersecurity standards include NIST, PCI DSS, ISO 27001, and CIS\_CSC, each offering best practices to achieve specific cybersecurity objectives.

### **2.7.1 NIST Cybersecurity Framework**

The NIST Cybersecurity Framework is a voluntary framework that provides organizations with guidance on managing and mitigating cybersecurity threats. While initially designed for critical infrastructure organizations in the US, the framework's flexibility allows its adoption by companies worldwide. The NIST standard enables the evaluation and improvement of existing cybersecurity strategies to enhance their performance (*NIST CYBERSECURITY FRAMEWORK*, 2018).

The need to align with particular cybersecurity frameworks arises from a variety of factors, including regulatory obligations, meeting industry regulator expectations, adherence to internal or external audit guidelines, fulfilling business objectives and customer needs, or simply to enhance an organization's cybersecurity strategy (Sabillon *et al.*, 2017). The NIST framework comprises three key components: the core, implementation tiers, and profiles. The core consists of functions, categories, subcategories, and informative references, offering recommended approaches to various aspects of cybersecurity. Its functions include identifying potential cyber risks, protecting against identified risks, detecting malicious activities, responding to threats, and recovering from breaches. The implementation tiers involve implementing risk management programs and processes to establish a robust cybersecurity framework, with categories including risk-informed, adaptive, partial, and repeatable. Profiles can be current or target, providing an overview of an organization's current cybersecurity system and its intended future state (*NIST CYBERSECURITY FRAMEWORK*, 2018).

### **2.7.2 PCI DSS**

The Payment Card Industry Data Security Standard (PCI DSS) was established by the PCI Security Standards Council to promote and maintain adequate security standards for the payment card industry (*PCI Security Standards Council*, 2023). It helps merchants and financial organizations understand and adopt security standards for technologies, policies, and processes involved in payment systems. By adhering to PCI DSS, organizations protect payment data from breaches and unauthorized access, ensuring the security of cardholders' information (*PCI Security Standards Council*, 2023).

### **2.7.3 ISO/IEC 27001**

ISO 27001, also known as IEC 27001, is an international standard that defines best practices for information security management systems (ISMS) (*ISO/IEC 27001, The Information Security*

(ISMS) Standard, 2022). Organizations certified with ISO 27001 demonstrate strict adherence to cybersecurity best practices and are considered to have effective data protection measures. This standard emphasizes the importance of risk management, a cornerstone of the information security management system (IT Governance, 2020). ISO 27001 programs focus on assessing risks associated with information security to establish robust controls.

#### **2.7.4 CIS Critical Security Controls**

The CIS Controls (CIS\_CSC) standards provide best practice guidelines for securing data on various electronic devices, including mobile devices, personal computers, and workstations (CIS Controls, 2023). These standards recommend rigorous configuration control and change management measures to prevent vulnerabilities from being exploited by cyberattacks (CIS Controls, 2023). By adhering to CIS\_CSC standards, organizations can protect their devices from exploitation and enhance their overall cybersecurity posture.

Incorporating these cybersecurity and control standards helps organizations establish comprehensive cybersecurity frameworks, safeguard sensitive information, and minimize the potential impact of cyber threats. Compliance with these standards demonstrates a commitment to implementing robust security measures and ensuring the confidentiality, integrity, and availability of data and systems.

#### **2.7.5 Cybersecurity Assurance Approaches**

Organizations In the field of cybersecurity, various approaches and frameworks have been developed to provide assurance and establish a strong security posture for organizations. These approaches aim to address the evolving threat landscape and the increasing complexity of technology systems. In this section, we will discuss some of the existing cybersecurity assurance approaches and their significance. One prominent cybersecurity assurance approach is the Cybersecurity Maturity Model Certification (CMMC). The CMMC was developed by the U.S. Department of Defence (DoD) to enhance the cybersecurity practices of organizations participating in the Defence industrial base. It is a unified standard for implementing cybersecurity across the Defence supply chain (Office of the Under Secretary of Defence for Acquisition & Sustainment, 2020). The CMMC framework provides a set of maturity levels and associated practices and processes that organizations must meet to achieve certification. It ensures that organizations handling sensitive defence information have the necessary cybersecurity controls in place to protect that information. Another widely recognized cybersecurity assurance approach is the Federal Risk and Authorization Management Program (FedRAMP) is a cybersecurity assurance program established by the U.S. federal government. FedRAMP provides a standardized approach to security assessment, authorization, and continuous monitoring of cloud services (Federal Risk and Authorization Management Program, 2021). It ensures that federal agencies have the necessary confidence in the security of the cloud services they use. FedRAMP sets stringent requirements for cloud service providers, including security controls, vulnerability management, incident response, and continuous monitoring. These cybersecurity assurance approaches play a critical role in establishing trust, enhancing cybersecurity practices, and providing assurance to stakeholders. By adopting and complying with these frameworks, organizations demonstrate their commitment to protecting sensitive information, managing risks, and mitigating cyber threats.

## 2.8 Data-Driven Threat Modelling Standards and Catalogues

Organizations often rely on traditional risk management practices to respond to attacks promptly. However, NIST 800-154 introduces dynamic data-centric threat modelling guidelines to proactively facilitate risk management processes, despite being in its early stages (Elahi *et al.*, 2021; Tatam *et al.*, 2021). This publication, titled "Introduction to Data-Centric System Threat Modelling," serves as a guide for threat modelling based on data and is currently in draft form since its release in 2016. The proposed model focuses on data within specific systems, such as stored data on laptops. The threat modelling approach is based on the concepts of Attack-side and Defence-side. The Attack-side discusses core terms like vulnerability, attack vector, threat, exploit, and attack, while the Defence-side addresses risk, security controls, and objectives. NIST's threat modelling approach consists of four steps: system and data identification, attack vector determination, security mitigation, and threat model analysis (Souppaya and Scarfone, 2016).

In understanding security-related adversaries, cybersecurity presents a complex pattern that demands multifaceted approaches. The Common Attack Pattern Enumeration and Classification (CAPEC) (MITRE CAPEC, 2023) organizes knowledge of adversary behaviour and focuses on specific use cases for application security ("CAPEC"). This model defines standard techniques and attributes used by attackers when exploiting vulnerabilities in cyber-enabled capabilities, such as clickjacking and session fixation. CAPEC is employed in application threat modelling and penetration testing to understand attacker perspectives, concepts and standardize countermeasures. Common Weakness Enumeration (CWE) (MITRE CWE, 2023) is a compilation of software and hardware security weaknesses aimed at addressing cybersecurity and everyday IT needs. CWE serves as a parameter for evaluating security tools while identifying, mitigating, and preventing attacks ("CWE").

The existing threat control standards, the International Organization for Standardization (ISO) and National Institute of Standards and Technology (NIST) are widely adopted for information security. ISO/IEC 27002 and NIST SP800-53 publications primarily focus on providing security controls. NIST is often referenced for threat mitigation due to several reasons. Firstly, NIST offers a more comprehensive set of controls, whereas ISO controls represent a subset of NIST controls, providing organizations with broader coverage for compliance requirements. Secondly, NIST's most recent version, NIST 800-53 revision 5, released in 2021, surpasses ISO 27002's 2013 version. Lastly, ISO charges for its publications, while NIST's publications are publicly available.

In summary, the mentioned works are essential for understanding and analysing threats in areas where existing threat modelling approaches do not strongly emphasize data. While most threat analysis focuses on assets, attacks, and threat modelling, NIST SP 800-154 addresses the need for a thorough understanding of data in the context of complete systems. Our research addresses these limitations by proposing a data-driven model that incorporates distinct types of data for comprehensive threat analysis.

## 2.9 Review and Discussion of Related Work

Cyber assurance has become a crucial element of any digital economy. The field has attracted the attention of several experts and scholars from all over the world. Most of the studies that are conducted on the subject revolve around the expected behaviours of cyber criminals and the anticipated trends in the associated attacks. According to leading industry research company (Gartner), they expect the trends that impact enterprises strategy that include the management of

threat exposure, validation of cybersecurity, security operation models, and composable security(*Top Strategic Cybersecurity Trends for 2023*, 2023). In Overall, most researchers focus their studies on the evolution of cyberattacks and the implications of such changes on global cyberspace.

Several works have discussed threat analysis extensively; most of these works rely on pre-existing models or standards that meet their research requirements, such as STRIDE and PASTA. The evolution of threat analysis models in cybersecurity, from foundational frameworks like Attack Tree, STRIDE, PASTA, and Kill Chain to more integrated and automated approaches, signifies the cybersecurity community's adaptive response to the complexities of modern cyber threats. This progression highlights a shift towards more sophisticated, dynamic, and context-aware methodologies that leverage advancements in technology and data analysis to enhance threat detection, assessment, and mitigation strategies. Modern techniques that rely on these models include hybrid threat modelling approaches that combine system-centric and attacker-centric perspectives for more comprehensive security analysis. For instance, Viswanathan and Prabhu (2021) propose a hybrid model integrating STRIDE and Attack Tree methodologies to identify threats during the software design phase, demonstrating its efficacy through a case study on a health centre management system (Viswanathan and Prabhu, 2021). Moreover, Straub (2020) explores the application of Blackboard Architecture to model attacks and defences, suggesting a generalized solution that accommodates various modelling techniques, including Kill Chain and STRIDE, for a more dynamic and flexible threat analysis framework (Straub, 2020).

Today, The integration of machine learning(ML) and AI into threat modelling represents another modern advancement. Techniques like text classification models are being developed to automate the mapping of threats within large datasets, enhancing the efficiency and accuracy of threat identification and prioritization processes. For instance, a study introduces a threat modelling for network intrusion detection system (NIDS) based on ML using STRIDE and Attack Tree approaches to identify the potential threats based on multiple levels, such as subtle perturbations inserted to the original inputs at inference time in order to evade the classifier detection or at training time to degrade its performance(Ali Alatwi and Morisset, 2022). Furthermore, this AI/ML is susceptible for unique vulnerabilities compared to traditional software systems. Hence, research provides an enhanced version of STRIDE to address this challenge by proposing an asset-centric approach(STRIDE-AI) for identifying threats to machine-learning-based systems(Mauri and Damiani, 2022). Conventionally, out of this literature review process, and despite of research advancement in threat modelling, there are four basic types of threat modelling, which include asset-centric, attacker-centric, software-centric, and threat-centric modelling tools (Yeboah-Ofori and Islam, 2019). As the names suggest, each of these techniques is intended to target particular components within a system. Since the most common approaches to threat modelling fail to consider the threats that are aimed at specific types of data, it remained difficult to deal with attackers whose malicious activities were intended to steal, destroy, or interfere with given data types (Shelupanov et al., 2019). Consequently, there is a pressing need to investigate new methodologies that prioritize data at the heart of threat analysis. Despite its critical importance in strengthening cybersecurity defences, data-centric threat analysis has not received sufficient attention from researchers.

Overall, the works discussed above offer workable methods for comprehending and analysing risks in fields where conventional threat modelling techniques do not concentrate a strong emphasis on data. Threat analysis typically places an emphasis on assets, tactics, actors, and threats.

Furthermore, business data has been the focus of standards like NIST SP 800-154 for systems that do not fully grasp all data from the entire system environment, such as control and management data. Our study addresses these limitations by offering a native data-driven threat analysis model that considers all relevant technological aspects of various forms of data present in the infrastructure at any given time during its lifetime. In addition to considering company operations and services as the initial point of comparison when analysing threats, the proposed methodology also considers shortcomings in earlier works. Hence, the proposed model in the research provides a superior value in several ways. First, the proposed model considers all the fundamental factors of data security, including actors, weaknesses, threats, controls, data, and infrastructure, which are then implemented using a conceptual model. Second, the approach considers data from three distinct levels of abstraction, such that the threat vector is analysed for data relating to management, business, and control. This approach ensures that the model offers full visibility and control over the data’s location at any part of its lifecycle by providing a layered representation of potential attack surfaces mapped to particular threat actors. Additionally, the model considers various steps in how the attack can escalate using our model. The use of three abstraction levels offers dynamism, enabling the model to adapt to the changes that characterize the field of cybersecurity. Since the model is dynamic, it is expected to support emerging concepts, like software-defined networking and cloud computing technology. The final property that makes this model superior to similar initiatives like NIST’s data-centric modelling system is that it is applicable and not limited to a particular system or data type i.e., business, or operational data.

In addition to the aforementioned points, the proposed data-driven threat analysis model provides a comprehensive threat analysis that goes beyond merely identifying threat vectors. It includes evaluating the effectiveness of the identified controls, which are crucial in supporting organizations to ensure their security assurance. By evaluating the identified controls, the model enables organizations to assess the adequacy of their security measures and determine whether they align with industry best practices and standards. This evaluation process involves analysing the capabilities and limitations of the controls in mitigating the identified threats. It provides organizations with valuable insights into the effectiveness of their existing security measures and helps identify potential gaps or areas for improvement. The inclusion of control evaluation within the data-driven threat analysis model not only enhances organizations' understanding of their security posture but also facilitates the development of effective risk mitigation strategies. By identifying areas where controls may be insufficient or ineffective, organizations can prioritize their resources and efforts to strengthen their security measures where they are most needed. Finally, to present a summary of the discussed works, Table 2.3 illustrates some threat modelling aspects of existing models.

Aspect	d-TM	Attack Tree	STRIDE	PASTA
<b>Focus</b>	Data	Attacker, Systems, Applications	Threats, Systems, Applications	Risk, Systems, Applications
<b>Data Abstraction Levels</b>	Three levels, Management, Control, Business	No data definition	A single type, i.e., business	A single type, i.e., business
<b>Threat Layers</b>	Agent, Network, Application, Compute, and Storage	Focuses on the paths an attacker might take, and does not specify layers.	Considers threats at different layers based on the type but is not structured	Assess threats within the context of the risk to business objectives and

			around specific layers.	technical environment.
<b>Process Automation and visualization</b>	Automated, and Comprehensive visual outputs, including color-coded tables, DFDs, and reports	Not inherently automated; emphasizes manual, risk-centric analysis with the option for visual representation.	Can be used in conjunction with automated tools; visual output can vary.	Not typically automated; visual output is a tree structure showing attack paths.
<b>Threat Mitigation</b>	Includes control evaluation as part of the process	Identifies potential attacks but does not typically include mitigation within the model.	Focuses on the identification of threat types with limited guidance on mitigation strategies.	Integrates threat analysis with risk management and mitigation planning.
<b>Adaptability to New Technologies</b>	Designed to adapt emerging tech like cloud and SDN	Traditional method; may not directly address emerging technologies without modification.	Flexible enough to be applied to new technology contexts.	Risk-centric approach adaptable to new technologies and business processes.
<b>Security Assurance</b>	Emphasizes security assurance by evaluating controls	Provides a structure for assessing potential attacks but is not directly tied to control evaluation.	Aims to provide assurance by addressing all six threat categories.	Designed to align threat modelling outcomes with business risk and security assurance requirements.
<b>References</b>	(Alwaheidi and Islam, 2022)	(Saini, Duan and Paruchuri, 2008)	(Hernan <i>et al.</i> , 2006)	(Ucedavélez and Morana, 2015)

Table 2.3. Overview of d-TM model to existing works

In conclusion, the continuous evolution of threat analysis models in cybersecurity reflects the field's dynamic nature and the ongoing efforts to develop more effective, efficient, and context-aware strategies to counteract sophisticated cyber threats. The integration of traditional models with modern technologies and methodologies marks a significant advancement in the cybersecurity domain, offering promising avenues for future research and application in safeguarding digital assets and infrastructures. The proposed data-driven threat analysis model addresses limitations in today and traditional threat modelling approaches by emphasizing the importance of data security and incorporating control evaluation. By considering all relevant technological aspects of various data forms and utilizing a layered representation of potential attack surfaces, the model provides a comprehensive understanding of the security landscape. Additionally, evaluating the effectiveness of identified controls enables organizations to assess their security measures, identify gaps, and prioritize resources for strengthening security. Overall, this holistic and dynamic approach supports organizations in enhancing their security posture, mitigating risks, and ensuring their security assurance.

## CHAPTER *THREE*: RESEARCH METHODOLOGY



## **3.1 Introduction**

This chapter presents the overall methodology used by this research. The research methodology adopts a combination of qualitative and quantitative methods to construct the study's content. The qualitative method involves an extensive literature review to explore existing knowledge on threat modelling and data-driven analysis techniques. This review helps identify gaps in the current state of the art and provides a foundation for the proposed model. The quantitative method, on the other hand, focuses on the analysis phase of the research and utilizes algorithmic approaches to simulate potential threats to which the data may be exposed at different instances.

## **3.2 Research Methodology**

Research methodologies serve as the backbone for conducting thorough and impactful studies. This section delves into the existing methodologies within three distinct research areas: literature review, model development, and evaluation. Furthermore, this section illustrates adopted research methods for addressing research questions and achieving objectives.

### **3.1 Existing Research Methodologies**

There are several methods used to conduct research, representing fundamental approaches in academic research across various domains. Each methodology offers a unique perspective and approach to exploring and understanding complex research questions, thereby contributing to the advancement of knowledge within their respective fields. The research considers three domains to investigate: Literature review process, development of an integrated models, and evaluation methods for validating research model.

#### **3.1.1 Literature Review**

There are several approaches to conduct this activity, literature review methodologies are instrumental in synthesizing existing knowledge, identifying gaps in current research, and setting a foundation for new inquiries. These methodologies enable scholars to critically assess and integrate findings from various studies, thereby contributing to the advancement of academic disciplines. Here, we discuss prevalent methodologies employed in conducting literature reviews, highlighting their academic significance. There are some common approaches including, narrative reviews, systematic reviews, and scoping review. Narrative reviews, often known as traditional literature reviews, involve a qualitative synthesis of a broad range of research related to a specific topic or question. These reviews provide a comprehensive overview, allowing scholars to trace the development of theories, methodologies, and findings within a field. Narrative reviews are particularly useful for areas where a formal meta-analysis might not be feasible due to the heterogeneity of studies. However, they may be susceptible to bias as the inclusion of studies is not always systematic (Green, Johnson and Adams, 2006). The Systematic Literature Review (SLR) which is a structured approach aimed at comprehensively collecting and critically analysing research evidence related to a specific research question or area of interest. Siddaway et al. (2019) emphasize that SLRs are characterized by their methodical and replicable methodologies, including a comprehensive search strategy to locate all relevant work, both published and unpublished, and a systematic integration and critique of the evidence to draw broad theoretical conclusions (Siddaway, Wood and Hedges, 2019). Similarly, Armstrong et al. (2011) highlight the importance of defining the scope of the research question in SLRs to clarify definitions and conceptualizations within the existing literature, thereby guiding the review process (Armstrong

*et al.*, 2011). Furthermore, Scoping reviews aim to map the key concepts, types of evidence, and gaps in a research area, particularly when the topic is complex or has not been comprehensively reviewed before. Unlike systematic reviews, scoping reviews tend to address broader questions and involve a more flexible approach to study selection and synthesis. This methodology is useful for clarifying working definitions and conceptual boundaries in emerging fields (Arksey & O'Malley, 2005). While, Recent advancements in literature review methodologies emphasize the importance of systematic and scoping reviews for synthesizing research findings comprehensively. Snyder (2019) outlines literature review as a critical research methodology, providing guidelines for conducting and evaluating literature review papers, ensuring thoroughness and rigor (Snyder, 2019). In conclusion, the choice of literature review methodology depends on the research question, the nature of the available literature, and the review's objectives. Each methodology offers distinct advantages and is suited to different types of inquiries, contributing to the richness and diversity of academic scholarship.

### **3.1.2 Model Development**

In the research context, research model development methodologies are fundamental tools that facilitate the systematic investigation of complex phenomena across various disciplines. These methodologies enable researchers to abstract, conceptualize, and operationalize the components and dynamics of their study domains, thereby providing a structured approach to inquiry. This discussion elucidates some of the prevalent methodologies in research model development, underscoring their academic relevance and application. Conceptual modelling stands out as a pivotal methodology in research model development, offering a blueprint for understanding and representing the abstract structures of a study domain. It involves the creation of conceptual frameworks that delineate the key concepts, constructs, and their interrelations within a research field. Robinson (2008) emphasizes the importance of a structured approach to conceptual modelling, which includes defining the problem domain, identifying objectives, and specifying the model's structure and content. This methodology is instrumental in ensuring that the research model is both comprehensive and aligned with the research objectives (Robinson, 2008). Ontological approaches in model development aim to ground the conceptual models in a well-defined and universally accepted set of entities and their relationships. This methodology is particularly valuable in ensuring semantic clarity and consistency across research models. Guarino and Welty (2009) advocate for the use of formal ontology as a foundation for conceptual modelling, facilitating the precise definition and classification of the entities and phenomena within a research domain. The ontological approach enhances the rigor and interoperability of research models, making them more understandable and reusable across different studies and disciplines (Guarino and Welty, 2009). While, Agent-based modelling (ABM) offers a distinctive approach to research model development by simulating the actions and interactions of autonomous agents within a system. This methodology is particularly suited to exploring the emergent behaviours and phenomena that arise from individual-level interactions. Bonabeau (2002) underscores the flexibility and adaptability of ABM in modelling complex adaptive systems, where the behaviour of the system as a whole cannot be easily deduced from the properties of individual agents. ABM facilitates the exploration of how local rules and behaviours lead to global patterns, making it a powerful tool for studying social, economic, and ecological systems (Bonabeau, 2002). In summary, research model development methodologies such as conceptual modelling, ontological approaches, and agent-based modelling offer diverse and powerful tools for academic inquiry. Each methodology provides a unique lens through which researchers can conceptualize,

operationalize, and analyse the intricate aspects of their study domains, contributing to the advancement of knowledge and understanding across various disciplines.

### **3.1.3 Model Evaluation**

Model evaluation is a crucial part of any research that aims to propose a new model to the knowledge. There are several methodologies to achieve this objective, such as Empirical investigation, and action research methodology. Empirical investigation in software engineering is a widely accepted validation method that evaluates proposed techniques through practical experience. It focuses on investigating the benefits and limitations of the technique (Aftab *et al.*, 2018). In contrast to other types of model evaluation methodologies, empirical investigation involves implementing client-oriented projects and analysing the results through empirical analysis (Breed and Verster, 2019). Furthermore, Empirical investigations, particularly through case studies, are a crucial methodology for evaluating theoretical models and hypotheses within real-world contexts. Case studies provide in-depth insights into specific instances, allowing researchers to explore the dynamics and complexities of phenomena. Linnenluecke, Marrone, and Singh (2020) detail methodological steps for conducting literature reviews and bibliometric analyses in a replicable and scientific manner, which can be applied to empirical case studies to enhance their analytical depth and reliability (Linnenluecke, Marrone and Singh, 2020). On the other hand, Action research methodology involves the identification of problems, collection of information, improvement of performance based on solutions, and evaluation of intervention results. The Simmons model is a comprehensive model for action research, which includes steps such as subject identification, data collection, planning, implementation, evaluation, and feedback (Aghdash *et al.*, 2021). Action research appeals to researchers and organizations seeking impact and utilization of scientific results in practice. It can involve customers through experiment systems and can be managed effectively (Staron, 2020). Action research is a methodology used to identify necessary changes within an organization, providing quality information for decision-making (Pracht, Toelle and Broaddus, 2022). It merges research principles and theories into practice, producing relevant research findings (Khan and Manzoor Rashid, 2022).

### **3.2 Research Methodology for d-TM**

This section presents how the chosen methodology is applied for d-TM. The research methodology encompasses Four stages, i.e., literature review, then model and tool development, lastly evaluation and conclusion. Each part contributes to the overall development and evaluation of the data-driven threat modelling and analysis approach, as illustrated in Figure 3.1. The initial stage narrates the gaps in the existing knowledge, and the second stage aims to propose a novel model and tool to address the gap in the knowledge. The third stage includes the evaluation of the proposed model and tool using multiple case studies. Lastly, the research conclusion and confirmation of the presented questions and objectives are addressed. Furthermore, the figure presents adopted methodologies for each stage, such as SLR, for literature review.

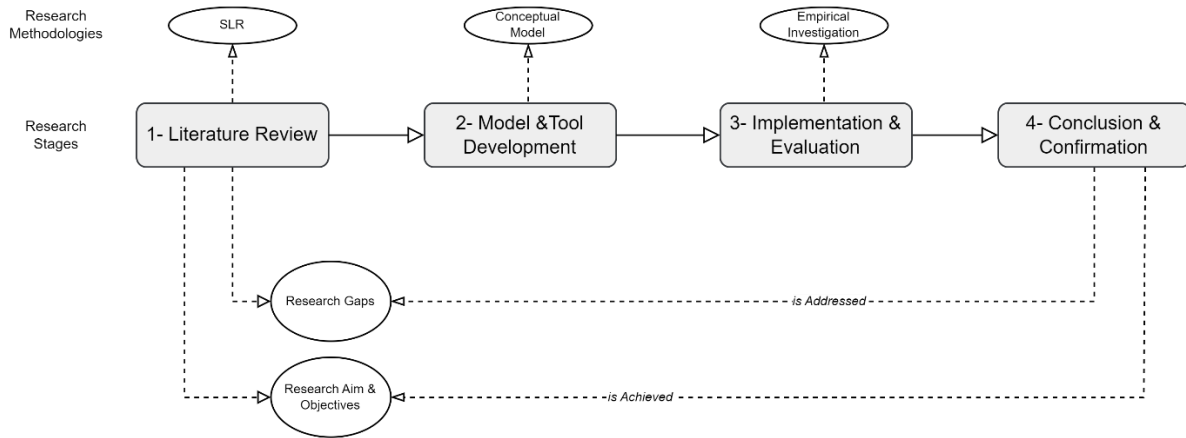


Figure 3.1. Methodology of the study

### 3.2.1 Stage 1. Literature Review

This initial stage intends to explore, analyse, and conclude the knowledge of existing approaches and gaps in the state-of-the-art domain knowledge. The research considers Systematic Literature Review (SLR) methodology for conducting the state-of-the-art review in threat modelling and analysis approaches is well-justified. It stands out as an optimal methodological choice for several reasons, particularly for its comprehensive and structured approach to literature review. Also, it aligns with the research's aim to systematically explore and synthesize domain knowledge, ensuring that the review is comprehensive, methodologically sound, and aligned with scholarly best practices. This activity is carried out through the exploration of pertinent databases, such as IEEE Xplore, UEL Library, CM Digital, Scopus, and Google Scholar. To compile articles that pertain to the subject of interest in the cybersecurity and threat modelling domain. The selection process is then executed and involves the utilization of keywords and filters to control the publication date ranging from (2018-2023) and the type of publications (Articles/Journals & Conferences) to identify the fundamental papers for subsequent analysis. The assessment stage includes the relevance of the subject, research title, type of research, date of publication, the content focus of the research, the research language, and author keywords. By applying these factors, the assessment resulted in high-quality and pertinent studies for analysis. Lastly, the analysis of the literature entails the scrutiny of patterns that arise among keywords, thereby unveiling trends, gaps, and challenges within the domain. To ensure a comprehensive review, specific criteria are set for searching and selecting relevant primary studies in the field of threat modelling and data-driven approaches for cybersecurity assurance. The research subject area and title should be within the context of cybersecurity threat modelling, with a preference for studies that highlight data-driven challenges and concerns. Peer-reviewed published journals are given priority as the material type, and a publication date restriction for the last five years is applied to ensure up-to-date information.

### 3.2.2 Stage 2. Model and Tool Development

This step focuses on the development of a novel data-driven threat modelling and analysis approach(d-TM). The model incorporates various activities and concepts to address challenges facing the industry in safeguarding their data efficiently. This stage includes the definition of the key components of the model, including data-levels, phases, attack layers, and actors. The

requirements analysis lays the foundation for the d-TM model, where the model should provide a holistic and comprehensive approach to data-centric threat modelling and analysis to bridge the gap in existing works. To achieve this activity, a conceptual modelling approach is adopted. The adoption of conceptual modelling in this research is fundamentally justified by its ability to simplify and organize complex subjects, making them more analysable. This methodology provides a clear framework for identifying critical components and their interactions within the study area, enhancing the depth and clarity of the research. In this stage, the development of the d-TM model is conducted by understanding and analysing the requirements and expectations of the innovative model and tool to address research gaps and questions identified in the previous stage. A comprehensive conceptual model for data-driven threat modelling has been developed based on the outcome of the requirement analysis process. The model includes critical elements such as actors, threats, data, infrastructure, weaknesses, controls, and cybersecurity assurance. The overall objective of this conceptual model is to provide a systematic approach to facilitate the understanding, analysis, and management of threats to an organization's cybersecurity posture. The integration of these critical elements into the model aims to provide a solid foundation for effective decision-making and security implementation.

As an integral part of this research stage, a specialized tool is developed to automate the threat analysis process using the data-driven threat modelling (d-TM) approach. This tool aims to streamline and enhance the efficiency of threat analysis by leveraging the power of data-driven techniques and incorporating key functionalities to assist decision-makers in their cybersecurity efforts. The d-TM tool is designed to provide a user-friendly interface that enables organizations to visualize their data assets, infrastructure, and associated threats. Through interactive data visualization techniques, decision-makers gain a clear and comprehensive understanding of the cybersecurity landscape within their organization. One of the essential features of the d-TM tool is its capability to identify weaknesses within the organization's data assets. By analysing various data sources, such as configuration files and descriptive information about data asset deployment, the tool can identify potential weaknesses that could be exploited by threat actors. Furthermore, the tool incorporates a methodology to prioritize identified threats. Through a comprehensive analysis process, threats are evaluated based on their potential impact on business continuity, enabling decision-makers to allocate appropriate resources and prioritize mitigation efforts effectively. The d-TM tool also aims to provide customizable reporting functionalities, allowing decision-makers to generate detailed reports on identified threats, weaknesses, and recommended controls for cybersecurity assurance. These reports serve as valuable resources for informed decision-making and facilitate communication among stakeholders. By providing decision-makers with a comprehensive cyber threat analysis platform, the d-TM tool empowers organizations to proactively identify, assess, and mitigate threats, thereby enhancing their cybersecurity assurance. Finally, the development of the d-TM tool represents a significant contribution to this research, as it enables organizations to leverage data-driven approaches for efficient and effective threat analysis. The tool's functionalities, including data visualization, weakness identification, and threat prioritization, empower decision-makers to make informed decisions and take proactive measures to protect their critical assets and ensure cybersecurity readiness.

### 3.2.3 Stage 3. Model and Tool Implementation and Evaluation

This stage aims to ensure that the model and supporting tool can accurately identify, analyse, and prioritize threats based on the organization's specific data context. The implementation and evaluation of the developed innovative d-TM model is a crucial step to assess its effectiveness and applicability. The evaluation of the d-TM model consists of multiple real-case scenarios that are carefully selected to represent diverse cybersecurity challenges faced by organizations, including supply-chain, healthcare, and service provider business sectors. These scenarios encompass various industries, data types, and threat landscapes, providing a comprehensive evaluation of the model's capabilities. In the context of the evaluation methodology, the research considers Empirical investigation methodology to conduct this activity. This methodology is one of the common approaches that used by scholars, and allows for the systematic collection, analysis, and interpretation of data directly derived from real-world observations or experiments, providing a robust foundation for testing hypotheses and theories. However, the use of empirical investigation methodology significantly strengthens the research's contribution to the field, offering insights that are both credible and applicable to practical contexts. This stage incorporated two steps, implementation and evaluation.

The implementation is the first activity of the research that involves bringing the proposed data-driven threat modelling approach (d-TM) into practice. To ensure the effectiveness of the d-TM model, processes are designed to handle the characteristics and complexities of various data types present within an organization. These processes incorporate data analysis, weakness recognition, threat identification and mitigation. The goal is to leverage the data-driven approach to provide a comprehensive understanding of potential threats and their impact on the organization's cybersecurity assurance. Throughout the implementation process, continuous testing and observation are employed to ensure the applicability and reliability of the d-TM model. Real-world scenarios are utilized to evaluate the performance and efficacy of the model in identifying and analysing threats across different organizational data. By successfully implementing the d-TM approach and refining the processes and methodologies, this research aims to provide organizations with a robust data-driven model and tool for enhancing their cybersecurity assurance. This activity incorporates data collection, which aims to gather primary data related to business operations and data assets for the evaluation of the d-TM Model and tool. A combination of interviews and surveys with business stakeholders, cybersecurity experts, infrastructure, and systems professionals from various organizations should be conducted. This primary data collection process aims to obtain valuable insights into current business operations, types of data, running digital services, and infrastructure supporting business operations.

- Interviews: Interviews are conducted by cybersecurity experts who possess in-depth knowledge and experience in threat modelling. These interviews provide an opportunity to engage in detailed discussions and gather qualitative information about their business, data, services running business operations and underlying infrastructure. Through these interviews, valuable insights are obtained regarding the organization's specific context, challenges, and best practices related to threat modelling.
- Surveys: Surveys are utilized as an effective means to automate and systematically collect information from a diverse range of stakeholders within the organization. The surveys are thoughtfully designed to capture data pertaining to various aspects of the organization's operations, data assets, priorities, and infrastructure. By engaging a wider audience from

different departments within the organization, surveys provide a comprehensive understanding of the organization's unique perspectives, practices, and requirements for effective threat modelling. The collected data through surveys support a more holistic view of the organization's cybersecurity landscape and contribute to the development of the d-TM Model.

The data collected through interviews, surveys plays a crucial role in informing and validating the d-TM Model - no personal data is collected. It provides a foundation of real-world insights, industry practices, and practical challenges, enabling the development of a robust and effective data-driven threat modelling approach. The data collection outcome involves the gathering of relevant data assets from organizations. These data assets may include configuration files, descriptive information about the data asset deployment, network diagrams, system documentation, and any other pertinent information related to the infrastructure and data landscape of the organization. These data assets serve as real-world examples for evaluating and validating the d-TM Model. In addition to the survey or interview collected data, the outcome of this activity could include supporting files, including asset configuration files, or descriptive files.

- **Configuration Files:** Configuration files from data assets are collected to gain a deeper understanding of the technical aspects and settings of the systems and applications. These files provide insights into how the data assets are configured and managed within the organization's infrastructure. By analysing these configuration files, potential vulnerabilities and weaknesses can be identified, contributing to the development of robust threat models.
- **Descriptive Files:** Witnessing the data asset configuration involves collecting descriptive information about the deployment of data assets. This can include details such as the location of data storage, access controls, encryption mechanisms, and data flow diagrams. By observing and documenting the data asset configuration, the d-TM Model can capture critical information about the organization's data landscape and identify potential areas of concern from a cybersecurity perspective.

The next step is the evaluation process involves applying the d-TM model to various use cases, each case scenario is used to examine its performance in identifying and analysing threats and validate its applicability and usability. This includes assessing the model's ability to recognize different attack vectors and highlight weaknesses within the organization's data assets. In addition to the technical aspects, the evaluation process also considers the practical applicability of the d-TM model within the organization's context. This involves assessing factors such as accessibility, usability, scalability, and ease of implementation. This process is conducted in collaboration with cybersecurity experts, practitioners, and relevant stakeholders who provide feedback and insights based on their experience and expertise. Their inputs contribute to refining and enhancing the d-TM model, ensuring its alignment with real-world cybersecurity challenges and requirements. By implementing and evaluating the d-TM model using multiple real-case scenarios, this research aims to demonstrate its effectiveness in addressing the unique threat landscape of organizations and providing actionable insights for enhancing cybersecurity assurance. The validation process provides evidence of the model's practical utility, reliability, and ability to assist organizations in making informed decisions regarding threat mitigation and risk management. Overall, the d-TM model validation serves as a critical step in establishing its credibility and ensuring its relevance and usefulness in real-world cybersecurity scenarios.

#### **3.2.4 Stage 4. Research Conclusion and Confirmation**

This stage is fundamental in establishing the research's academic consistency and relevance. It thoroughly synthesizes the findings, ensuring they are directly linked to the research questions and objectives. Through a detailed comparative analysis with the d-TM model and existing literature, this section validates the research outcomes, highlighting both alignment and novel contributions. Also, This stage delineates the d-TM model and tool limitations, offering a transparent overview of potential biases and constraints. This critical self-assessment paves the way for future research, suggesting areas for further exploration and improvement. Ultimately, this section solidifies the research's position within the academic landscape, affirming its contribution to the field and setting the stage for ongoing scholarly dialogue.

### **3.3 Summary**

The research methodology for developing the data-driven threat modelling (d-TM) approach encompasses four major stages. Firstly, the literature review provides a comprehensive understanding of industry challenges and limitations in cybersecurity assurance, focusing on data-driven threat modelling. Secondly, the model and tool development formulates a framework that incorporates key elements such as actors, threats, data, infrastructure, controls, weaknesses, and cybersecurity assurance. Furthermore, a tool is developed to automate the threat analysis process, incorporating functionalities for data visualization, weakness identification, and threat prioritization. Model implementation and evaluation is the third stage which involves the implementation of the d-TM processes to analyse threats based on distinct types of organizational data. The effectiveness and applicability of the model are validated through real case scenarios, ensuring its accuracy in identifying and analysing threats. Lastly, the conclusion and confirmation stage, which concludes research findings, alignment to the research's objectives and questions, contribution, limitation and future works. These stages collectively contribute to the research goal of providing organizations with an effective data-driven approach for cybersecurity assurance, ultimately enhancing their resilience against cyber threats.



# CHAPTER *FOUR*: DATA-DRIVEN THREAT MODELLING (d-TM)

## 4.1 Introduction

This chapter presents one of the main contributions of the thesis, the data-driven threat model (d-TM), a novel approach that aims to analyse threats across the lifecycle of data comprehensively. The model provides organizations with valuable insights to guide essential management decisions for strengthening overall security capability and assurance. By bridging the realms of business operations and technology, the d-TM explores the complexities of organizational data in a systematic manner. The primary objective of this model is to empower business stakeholders with the requisite understanding of data and associated threats to make informed decisions for ensuring the continuous operation and security of their business.

## 4.2 d-TM Requirements

To achieve the intended objectives of developing a new threat analysis approach, certain requirements must be considered when developing the proposed d-TM model. These requirements are essential to ensure the functionality and efficiency of the model. The development of the d-TM model requirements was meticulously informed by the overarching research objectives, aimed at innovating a data-driven threat modelling approach to enhance cybersecurity assurance within organizations. Each requirement was crafted to align closely with these objectives, ensuring a targeted and functional model that addresses the specific needs identified through initial research phases. The main model requirements are outlined as follows:

- **Requirement 1:** The d-TM model shall enable users to systematically catalogue and assess assets and services within an organizational context, serving as a foundational step for subsequent threat analysis.
  - Linking to Objectives(Traceability): Objective.1 and Objective.2
  - Measuring Factor: the total number of assets and services identified within the organizational context.
- **Requirement 2:** The d-TM model shall delineate and employ a conceptual framework that encapsulates key concepts pertinent to threat analysis and management, ensuring a structured approach to identifying and addressing potential threats.
  - Linking to Objectives(Traceability): Objective.2 and Objective.3
  - Measuring Factor: through the completeness and applicability of the conceptual model in capturing data-related threats in real-world case scenarios.
- **Requirement 3:** The d-TM model shall possess the capability to anticipate and map out potential weakness and their corresponding threats through a comprehensive analysis, leveraging the defined data abstraction and phases to ensure a thorough threat landscape overview.
  - Linking to Objectives(Traceability): Objective.2 and Objective.3
  - Measuring Factor: by the model's ability to systematically uncover and articulate potential weaknesses and associated threats using real-word case scenarios.

- **Requirement 4:** The d-TM model shall integrate and utilize established security frameworks and knowledge bases, such as MITRE's Common Weakness Enumeration (CWE) and Common Attack Pattern Enumeration and Classification (CAPEC), to enrich and validate the threat analysis process.
  - Linking to Objectives(Traceability): Objective.2
  - Measuring Factor: by the integration and effective utilization of established security frameworks like CWE and CAPEC in threat analysis.
- **Requirement 5:** The d-TM model shall be adept at identifying and evaluating threats across various layers of organizational infrastructure and the relevant actors, utilizing the defined data abstraction and phases to facilitate a comprehensive threat analysis.
  - Linking to Objectives(Traceability): Objective.2 and Objective.3
  - Measuring Factor: by the model's capability to define and adopt various threat layers and actors, contributing to a nuanced threat landscape overview.
- **Requirement 6:** The d-TM model shall systematically identify potential mitigation strategies and select the most suitable controls for addressing identified threats, ensuring the effectiveness and relevance of the response measures.
  - Linking to Objectives(Traceability): Objective.2 and Objective.3
  - Measuring Factor: by the model's ability to propose relevant and practical controls for threat mitigation.
- **Requirement 7:** The d-TM model shall assess and verify the proposed controls, ensuring they provide the intended security assurance against identified threats.
  - Linking to Objectives(Traceability): Objective.2 and Objective.3
  - Measuring Factor: by the model's ability to evaluate and confirm the security assurance provided by the selected controls.
- **Requirement 8:** The d-TM approach shall be supported by an automated tool that streamlines the threat analysis process, enabling efficient and accurate derivation of analysis details and recommended controls.
  - Linking to Objectives(Traceability): Objective.2, Objective.3 and Objective.4
  - Measuring Factor: by the tool's ability to automate the threat analysis activities and mitigation for enhancing the model's applicability and efficiency.
- **Requirement 9:** The d-TM tool shall offer visualization capabilities such as DFD to depict all underlying services and assets, enhancing the understanding and analysis of the threat landscape from a data abstraction and phases perspective.
  - Linking to Objectives(Traceability): Objective.1, Objective.4

- Measuring Factor: by the tool's visualization capabilities, specifically its ability to accurately depict services and its underlying assets interconnectivity using DFD in the context of threat analysis.
- **Requirement 10:** The d-TM tool shall visualize identified weaknesses, threats and their criticality to business, utilizing the defined data abstraction and phases to present a holistic threat analysis.
  - Linking to Objectives(Traceability): Objective.2 and Objective.4
  - Measuring Factor: by the tool's ability to visually represent various identified weaknesses, potential threats and their criticality, facilitating informed decision-making.
- **Requirement 11:** The tool shall automatically assess and quantify the threat level for each identified threat, integrating with open intelligence and common security knowledge bases for a comprehensive threat context.
  - Linking to Objectives(Traceability): Objective.2 and Objective.4
  - Measuring Factor: by the relevance of the automated threat level calculations and their alignment with security open intelligence.
- **Requirement 12:** The tool shall autonomously evaluate and report the assurance level of the identified mitigation strategies, ensuring that the recommended controls effectively address the identified threats.
  - Linking to Objectives(Traceability): Objective.2 and Objective.4
  - Measuring Factor: Measured by the tool's capability to automatically determine and report the assurance levels of identified mitigation strategies, ensuring their effectiveness.

### 4.3 d-TM Fundamental Pillars

This section presents the main essential concepts used for the proposed d-TM. The proposed d-TM is based on a number of pillars, which provide the foundation for developing the threat assessment and management approach. Figure 4.1 depicts the d-TM pillars from a holistic perspective from the data level and phases for common security knowledge through threat layers and threat actors. The research discussed a number of pillars; that equip the d-TM model with the required tools that advance the threat analysis process and address the research objectives as well as the gap to existing works. As a result of the LR for academic and industry reports in threat analysis and modelling, the data-driven approach is overlooked(Pillar 1). Furthermore, developing data-driven approach necessitates the need for supporting concepts, such as data threat actors (Pillar 3), data threat layers – attack surfaces(Pillar 2), and lastly, to support the research objective in automating d-TM process, data threats and underlying weaknesses knowledgebase (Pillar 4) is provided to empower the threat assessment process with up-to-data intelligence and consistency in weakness and threat definition. However, The d-TM Pillars consider Data varieties, and their existence in organizational technology. In addition to, the potential threat actor and up-to-date weakness, controls, and threat-related, surface and knowledgebase. The proposed d-TM is

developed based on Four pillars(Data levels, Threat Actors, Threat Layers, and Common Knowledgebase), which provide the foundation for threat assessment and management approach to d-TM.

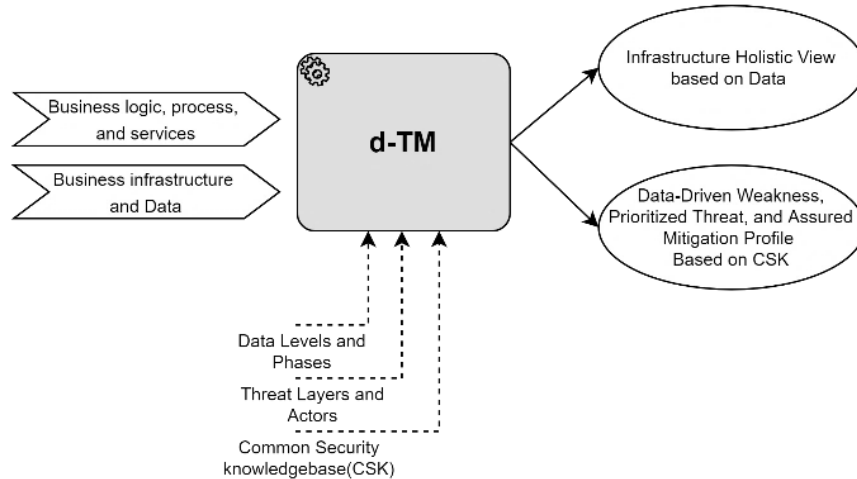


Figure 4.1. A high-level overview of d-TM.

### ***Pillar 1: DATA LEVELS AND PHASES***

The d-TM model acknowledges the significance of comprehending distinct types of data throughout its lifecycle. It categorizes data into three abstraction levels: management, control, and business(Alwaheidi and Islam, 2022). Each level encompasses three distinct phases of data: at rest, in process, or in transit. These abstraction levels hold equal importance, as an attack can directly or indirectly impact the business through any of these levels. The rationale behind incorporating abstraction levels and phases in d-TM is to ensure the security of data regardless of its location or status within the digital infrastructure. Furthermore, the risk is shared among these levels, where compromise at one layer could potentially escalate to another. For instance, if an attacker manages to steal authentication credentials (management data) of a particular asset, it can gain unauthorized access to the asset and compromise its functionality (control data), leading to the unauthorized transfer of business data. Figure 4.2 depicts a visual representation of the data levels within a digital infrastructure from the d-TM perspective, illustrating their interrelationships and a holistic view of how data flows and is processed within an organization's technological ecosystem.

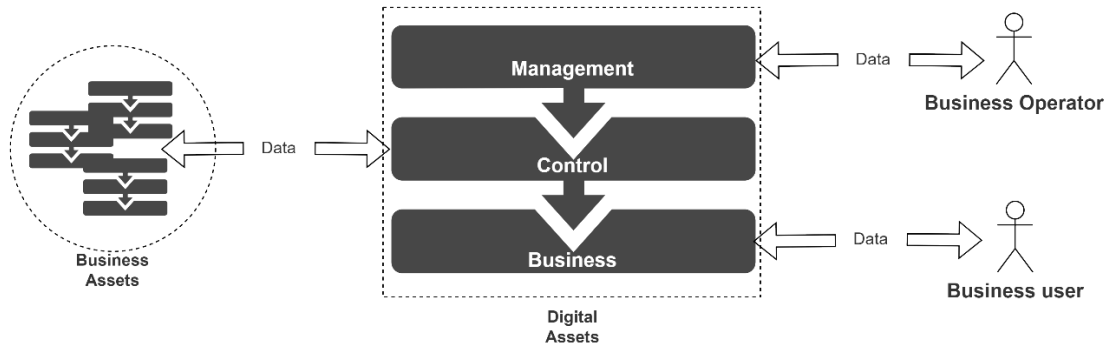


Figure 4.2. Data-levels at digital systems

The Figure illustrates a sequential relationship between data levels, notably from management to business. This sequence begins with the management data, followed by control data, and culminates with business data. To illustrate, consider the example of a newly acquired digital asset, such as a network router, which initially lacks any data or configuration. The initial step involves an administrator, referred to as the "business operator," accessing the router through management data-level, which includes credentials like the administrator's username, password, and IP address. Following this, the administrator configures the device to support its functionality by setting up dependent assets or services, such as DHCP server, DNS server, and APIs calls. The data associated with these configurations is termed as control data-level. Finally, once the device is configured and connected to the requisite services. In our case a network router, is capable of handling and processing business data-level, which is generated by business users. In summary, the lifecycle of data within any digital asset progresses from management to control, and ultimately to the stage where the asset is prepared to store, process, or transmit business data. This progression underscores the hierarchical nature of data handling within digital systems, emphasizing the foundational role of management data in enabling the subsequent layers of control and business data. The d-TM data level and phases are identified as follows:

- **Level 1. Management Data:**

This level represents data associated with system administration, including identity and access management. It involves administrators initiating data to access systems for various administrative activities (Admin-to-System). This data may include authentication details, authorization protocols, and access privileges. Ensuring the security of infrastructure management access is crucial, as compromising this layer is often the primary objective of attackers (Harris *et al.*, 2019). For example, stealing login credentials grants attackers superior control over computing or network devices. Management data is utilized for authentication and authorization purposes to access system functions (control data). The access can be user-based or system-based, employing various mechanisms such as SSH, Telnet, FTP, HTTP/s, Netconf, etc. Management data is vulnerable to attacks like brute force attacks, privilege escalation, session hijacking, and more.

- **Level 2. Control Data:**

This level encompasses data related to system functionalities and the exchange of data between systems to support business operations (System-to-System). Control data may include routing information, statistical updates, application inquiries, or configuration updates. It facilitates the sharing of data between systems, such as business-related information. For instance, applications

utilize Application Programmable Interfaces (APIs) for system-to-system information exchange, and network routers rely on protocols like Border Gateway Protocol (BGP) to exchange routing information. Control data is susceptible to various threats, depending on its presentation in the environment, whether local or remote. Insecure practices related to control data can result in significant network compromises. For instance, the absence of secure configurations for controlling information exchange between systems can introduce rogue information, leading to the manipulation or exfiltration of business-related data. Control data is vulnerable to threats like Man-in-The-Middle (MITM) attacks.

- **Level 3. Business Data:**

This level encompasses data directly related to an organization's business services (User-to-System). Business data is critical and runs on software systems that can be compromised if an attacker gains unauthorized access to control or management data. For example, in a network device, business data resides in the data plane, which receives instructions from the control plane regarding actions to be applied to specific traffic at specific times. If the control plane is compromised, the business data could be redirected to the attacker's system. Business data is susceptible to attacks that can disrupt business continuity, such as Distributed Denial-of-Service (DDoS) attacks. Figure 4.2 provides a high-level representation of the data levels within a digital infrastructure.

The understanding of data levels and phases within an infrastructure asset is succinctly summarized in Figure 4.3. It is important to note that data, as operated by assets, relies on three key components: the processor, memory, and input/output. At any given time, data can be in various states aka phases in the d-TM model, corresponding to different data levels. The figure is leveraged by colore coding style to show clarity, such as Red for data in transit, Green for data at rest, while Yellow represent data in use. For instance, if business data is at rest(green), that mean data is stored in asset memory(green), which is not processed by asset processor (yellow) nor transmitted over input/output interface(red), i.e., network interface.

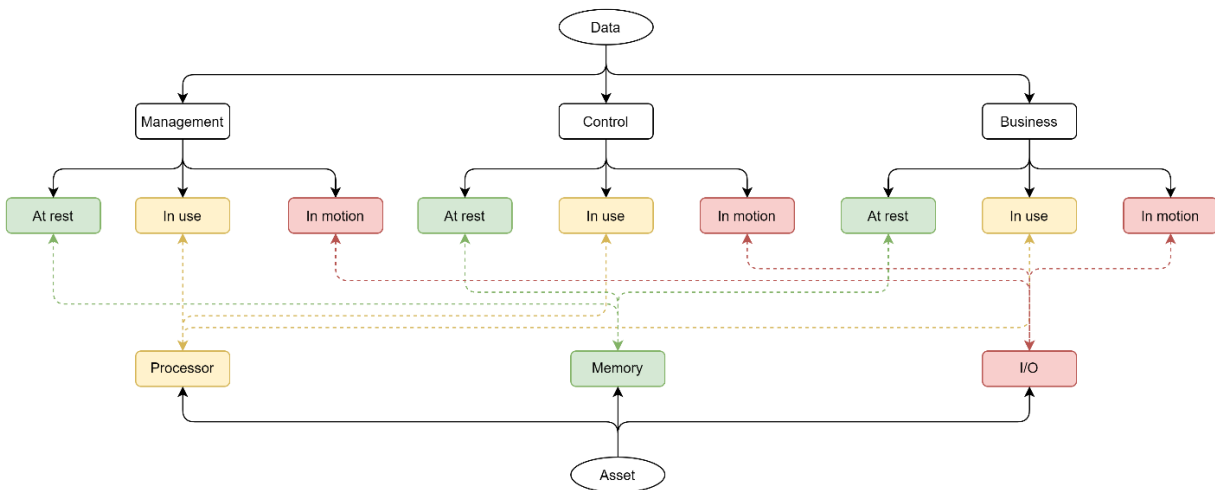


Figure 4.3. Data levels and phases

The data phase plays a pivotal role in demonstrating the state of the data at each level, namely at rest, in process, and in transit(Ullah *et al.*, 2018). Regardless of the phase, data must be protected at all levels within its lifecycle. Leading organizations, as highlighted by Westervelt (2020), recognize the importance of evaluating existing security controls to safeguard data at rest, in process, and in transit.

- **Phase 1** of the data lifecycle refers to **data at rest**, representing its initial and final state. During this phase, data is created and stored in local or remote storage, awaiting further processing or transmission. Data at rest is particularly vulnerable to cyber threats and unauthorized access. Safeguarding data at rest is crucial to prevent sensitive information from being leaked or stolen. Effective measures to protect data at rest include implementing robust network security, employing strong encryption techniques, continuously monitoring networks and infrastructure for suspicious activities, and diligently identifying and addressing potential vulnerabilities within the system and applications.
- **Phase 2** of the data lifecycle involves **data in process**, which refers to temporary data stored in memory and utilized during the execution of an application. While not inherently suspicious, this data can be manipulated by rogue services, indicating a potential security breach. Unauthorized access to system resources, employing advanced persistence threats like Stuxnet (Baker *et al.*, 2011), DuQu, Flame (Bencsáth *et al.*, 2012), or side-channel attacks (Abdulghani *et al.*, 2019) can compromise the integrity of in-process data. As the volume of data grows and more platforms access it, ensuring the security of in-process data becomes increasingly vital.
- **Phase 3** of the data lifecycle pertains to **data in transit**, which encompasses data that traverses between systems via computer networks. Data in transit is highly susceptible to security risks, particularly when transmitted over insecure channels or through application programming interfaces (APIs) that enable inter-application communication. Safeguarding data in transit is of utmost importance, irrespective of the growing regulatory focus on data protection. A data breach during this phase can have severe repercussions for a business, including the exposure of sensitive data, reputation damage, and financial penalties(Berecki, 2019).

In summary, the consideration of data levels and phases in the d-TM model enables organizations to adopt a comprehensive approach to threat analysis. By recognizing the significance of protecting data at rest, in process, and in transit, businesses can fortify their defences and mitigate potential risks associated with each phase of the data lifecycle.



## ***Pillar 2: THREAT LAYERS***

The d-TM (Data-Driven Threat Model) offers a threat analysis approach based on a generalized model of information technology organizational architecture. Organizations may adopt various technologies that suit their specific needs, such as networks, computers, and more. However, as technology continually evolves to offer greater functionality, flexibility, and business support, the threats and weaknesses also vary based on the chosen technology. Therefore, the threat analysis technique should be adaptable enough to handle diverse technology types. To address this, the proposed technique is built on a tier-based approach, enabling organizations to assess threats to data at any point in the infrastructure, regardless of data location.

In the d-TM model, each layer is designed to be intricately linked to organizational IT roles and response capabilities. For example:

- The Application Layer involves developers.
- The Compute Layer is managed by system administrators.
- The Network Layer is handled by the network team. And so on.

d-TM identifies five layers, each of which represents a potential attack surface. These layers illustrate the path of data flow from users to data stores and vice versa, and they are interconnected. Any vulnerability in these layers can lead to business disruption and data compromise.

- **Agent Layer:** This layer provides insight into the tools used by d-TM actors to access data or services, such as web browsers. Compromised or vulnerable web browsers pose significant threats to organizational data. Sensitive information like session details, encryption keys, or saved credentials could be exposed through rouge software/plugins or network attacks like Man-in-The-Middle (MITM) attacks. Securing the user agent is crucial, especially if the user has administrative access, as data compromise at this level could lead to a system takeover through stolen admin credentials.
- **Network Layer:** This layer identifies devices that interact with data before it reaches the business service or technology. It includes physical or virtual routers, switches, or load balancers. Business data is in transit at this stage and could be impacted by configuration manipulations that lead to data leakage. Attacks may come from external actors or even internal system admins due to misconfigurations. The network layer is eventually connected to the compute layer, which hosts the business application.
- **Compute Layer:** The compute layer represents the platform, software, or operating system that hosts the business application. It can involve virtual machines or container technology. Each computing technology requires individual assessment to ensure security. Manipulation in hosting OS services could disrupt all installed applications, guest VMs or containers. Once data is received by the compute layer, it proceeds to the application layer for processing.
- **Application Layer:** At this stage, data moves from transit to the process phase. The application layer is exposed to various threats due to its direct interface with internal and external networks. Unlike the compute or network layer, which does not interact directly with business users, the application layer is more susceptible to potential attacks.

- **Storage Layer:** This layer represents the final destination for data, where it is stored from the process phase to the at-rest phase. Storage can be local (attached disks) or remote (network storage). When data is stored in remote storage, threats during data transition or while being processed in network storage exist. Organizations must consider data location and security at every stage of the infrastructure and data lifecycle.

The d-TM approach allows organizations to adopt a holistic view of their data-driven threat landscape, enabling them to assess vulnerabilities and implement effective security measures at each layer of their infrastructure. By understanding and securing each layer, businesses can better protect their valuable data assets and maintain the continuity of their operations in the face of evolving threats.

### ***Pillar 3. THREAT ACTORS***

In the context of d-TM (Data-Driven Threat Model), the term "actor" refers to any human or machine attempting to gain access to an organization's digital resources. These actors may be authorized or unauthorized, and their intentions can range from benign to malicious. Furthermore, in the context of data abstraction levels, where data is processed, stored, or transmitted by an asset, we identify three principal actors that interact with the asset: the business user, the business operator, and the business-relevant system. Each actor engages with the asset in distinct ways and possesses the potential to influence the asset's functionality, thereby posing potential threats to data. These actors are delineated based on their interaction modalities with the asset, which range from direct operational control to indirect system interactions. Understanding the roles and potential impact of these actors is crucial for assessing and mitigating risks associated with data within an organizational framework. The approach considers four types of actors:

- ***Business-User:*** This type represents any legitimate human user accessing the organization's resources with the aim of benefiting from specific services.
- ***Business-Operator:*** Business Operators are legitimate human users who have access to the organization's resources for administrative tasks, such as updating, maintaining, or troubleshooting.
- ***Business-System:*** Business-Systems include any services, processes, or technologies that are legitimately connected to, or capable of connecting to, the organization's resources to support business or administrative functions. Examples include Internet of Things (IoT) devices.
- ***Threat-Actor:*** Threat-Actors represent any of the three previously mentioned actors (user, operator, or system) with the intent of abusing or disrupting business operations or gaining illegitimate access to sensitive data.

In the d-TM approach, actors play a significant role in the threat analysis process, whether they are internal or external to the business. Depending on the actor's characteristics and intentions, the company can identify potential attack scenarios that could disrupt the business. For example, if a threat actor manages to acquire the privileges of a business asset operator with access to critical organizational assets, the risk level could be significant, potentially leading to asset takeover and data infiltration. On the other hand, if the threat actor only possesses the access rights of a regular business user, the real threat may not be as severe.

The core concepts in the threat analysis technique of d-TM involve threat layers and actors in relation to data levels. Figure 4.4 illustrates a comprehensive understanding of these three concepts, providing a valuable framework for organizations to assess their threat landscape, identify potential vulnerabilities, and implement appropriate security measures at each data level and actor category. By incorporating the actor perspective into the analysis, d-TM enables organizations to develop targeted and effective strategies for safeguarding their digital resources and mitigating potential risks posed by various actors in their environment.

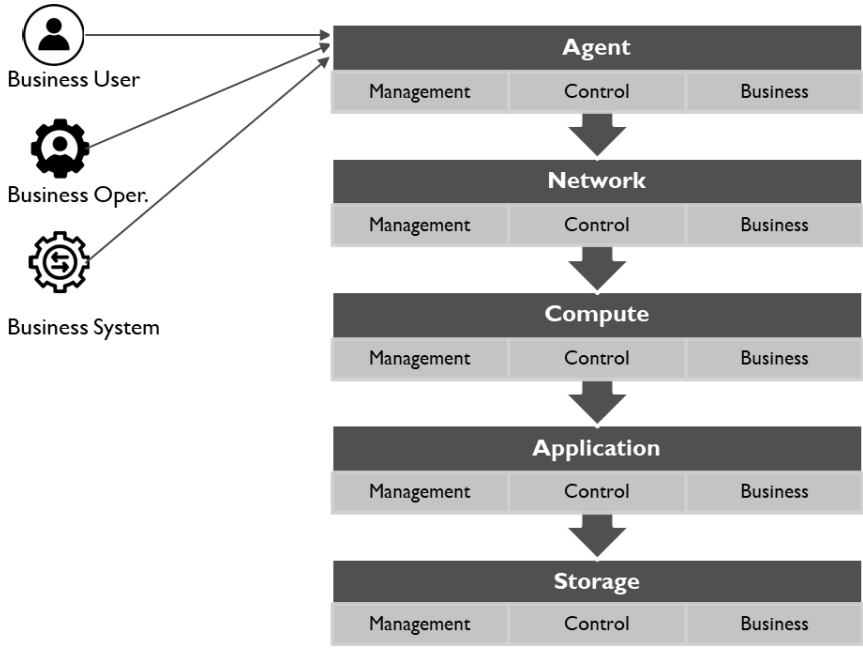


Figure 4.4. d-TM Data levels, layers and actor representation

**Pillar 4. COMMON SECURITY KNOWLEDGE BASE(Kb)**

In the context of d-TM (Data-Driven Threat Model), conducting a thorough threat analysis requires the utilization of various techniques. d-TM advocates incorporating three widely used knowledge bases: MITRE CWE (Common Weakness Enumeration) for weaknesses, MITRE CAPEC (Common Attack Pattern Enumeration and Classification) for threats, and NIST SP 800-53 for threat controls.

CAPEC serves as a valuable resource for organizing knowledge about adversary behaviour, focusing on specific uses for system and application security. It outlines typical approaches and properties that attackers employ when exploiting security weaknesses in cyber-enabled

capabilities. Examples include techniques like clickjacking and session fixation. By leveraging CAPEC, d-TM can better understand attacker perspectives, standardize countermeasures, and regulate security advancements to combat these attacks effectively.

Similarly, the Common Weakness Enumeration (CWE) knowledge base aids d-TM in creating a collection of security weaknesses in software and hardware. CWE provides essential information to comprehend the nature of flaws within a system. By utilizing CWE, d-TM gains the ability to govern recognized weaknesses and establish knowledge boundaries. Threats and weaknesses can then be connected, classified, and compared for any given system using CAPEC and CWE IDs, while also obtaining information on impact, mitigation, and associated dependencies. The last knowledge base is NIST SP 800-53 publications which are also embraced by d-TM, offering a comprehensive reference for a diverse set of controls to enhance the organization's cybersecurity posture.

- **Kb1. CAPEC**

In adopting CAPEC, this knowledgebase identifies and categorizes attacks based on six domains: Software, Hardware, Communication, Supply Chain, Social Engineering, and Physical Security. Each domain includes a list of attacks with relevant information such as related weaknesses, execution flows, prerequisites, and consequences. d-TM considers all these domains when working with organizations and data that may be linked to them. Furthermore, the attacks within each domain are categorized into three types of levels: Meta, Standard, and Detailed attack patterns. Meta-level serves as a top-level category, followed by Standard and Detailed patterns that offer specific information about threats.

While CAPEC focuses on categorizing threats based on its domains, d-TM expands its scope by considering the organization's infrastructure, including agents, networks, computes, applications, and storage, as potential targets. Table 4.1 represents the mapping of d-TM to CAPEC attack domains, considering the applicable CAPEC domains within the d-TM model. For example, "Compute," as one of d-TM's infrastructure elements, may be threatened by attacks related to Software, Hardware, Communication, Supply Chain, and Physical Security. However, Social Engineering attacks are not applicable to Compute, as they target people rather than the infrastructure.

d-TM Asset Layers	CAPEC Domains					
	Software	Hardware	Communication	Supply Chain	Social Engineering	Physical Security
Agent	Ö		Ö	Ö	Ö	
Network	Ö	Ö	Ö	Ö		Ö
Compute	Ö	Ö	Ö	Ö		Ö
Application	Ö		Ö	Ö		
Storage	Ö	Ö	Ö	Ö		Ö

Table 4.1. Mapping CAPEC domains to d-TM infrastructure categories

- **Kb2. CWE**

In the context of d-TM (Data-Driven Threat Model), it is crucial to consider the overall weaknesses within the infrastructure and system when analysing threats. There are two main methods widely

used for determining weaknesses: CWE (Common Weakness Enumeration) and CVE (Common Vulnerabilities and Exposures). d-TM advocates using CWE over CVE for several reasons:

- ***Easy understanding for non-experts:***

CWE is more easily digestible for both technical and non-technical security practitioners, including developers, system administrators, trainers, and business management. On the other hand, CVE focuses on specific codes, versions, and products, which can be complex for non-experts to comprehend. CWE allows organizations to understand their specific weaknesses rather than only common vulnerabilities.

- ***Lack of exploitation evidence by CVE:***

CVE provides limited exploitation information about vulnerabilities. CWE, on the other hand, provides the root cause and necessary traces behind vulnerabilities, enabling proactive investigation and prediction of vulnerabilities before exploitation occurs.

- ***Focus on vendor-specific products:***

CWE offers a wide range of error coverage compared to vulnerabilities that are specific to particular products and versions. Stakeholders in a system might only focus on publicly announced vulnerabilities without addressing the underlying weaknesses that could lead to such mistakes. CWE is widely adopted by major multinational tech companies, making it a comprehensive and industry-supported choice.

- ***Lack of proactive analysis of security vulnerabilities:***

Assessing weaknesses using CWE is a proactive approach to identifying errors in a system during early development stages or afterwards. On the other hand, CVE focuses on vulnerabilities in already-built commercial or open-source software. d-TM emphasizes CWE to provide organizations with oversight of potential weaknesses that could result from human or software errors at any stage, including operation and early development.

- ***Kb3. NIST SP 800-53 r5***

Regarding security controls, d-TM looks to NIST SP 800-53 revision 5 as a reference for threat mitigation. NIST, along with ISO, is widely adopted for providing information security. NIST SP 800-53 offers a more comprehensive set of controls compared to ISO/EC 27002, with ISO controls being a subset of NIST controls. Therefore, d-TM adopts NIST to provide organizations with broader coverage for compliance requirements. Moreover, NIST SP 800-53 revision 5 is the most recent version, released in 2021, compared to the 2013 version of ISO 27002. Additionally, NIST publications are available to the public, while ISO charges for its publications.

## **4.4 Conceptual View of d-TM**

The conceptual view of d-TM is based on a holistic understanding of data and its surrounding environment. The model incorporates several relevant concepts for threat analysis, considering data from three different abstraction levels. It offers a comprehensive analysis of each layer

involved in the data lifecycle and examines attacks by defining multiple attack surfaces and potential threat actors aligned with modern organizational operational layers.

The proposed d-TM model, depicted in Figure 4.5 that combines all concepts to support threat assessment and management is based on specific conceptual relationships and clearly defined attributes. Threat modelling and definition within this model consider the specific organizational data. The foundational concepts of d-TM are explained in detail in previous sections; however, it is summarized in this section as follows:

**Data:** Recognizing the importance of data in the threat modelling process, the model places significant emphasis on understanding and analysing distinct types of data within the organization. The data element is further refined to focus on three abstraction levels of data: management, control, and business. Additionally, it considers the divergent phases of data: at rest, in process, and in transit. This granularity allows for a comprehensive analysis of data throughout its lifecycle and provides a more detailed understanding of potential vulnerabilities and threats.

**Actor:** Understanding the actors within the organizational context is paramount in threat analysis. It allows for the identification of vulnerabilities and potential weaknesses that could be exploited by threat actors. Moreover, recognizing the motivations and behaviours of these actors is essential in developing effective security controls and responses. The conceptual model considers three actors within the organizational context, as illustrated in the d-TM foundation pillars. The actor could be a business user, operator, or system; any of these actors could be a potential threat actor. However, the Actor represents the users of the data, which can be business or operation oriented.

**Asset:** Represents any hardware or software utilized by the organization to access or operate underlying business services. Assets are categorized into five types: agent, network, compute, application, and storage. Agents assist users or administrators in accessing business assets, while the network provides visibility on devices facilitating access to organizational services. Compute and Storage assets define application hosting environments and data storage locations, respectively.

**Threats:** Refers to the potential of performing malicious acts that could harm the organization's infrastructure or data, exploiting weaknesses within the system. These weaknesses could be related to code, configuration, or architecture flaws in the system. The model defines several types of threats based on common security knowledgebase that organizations may encounter, including both known and emerging threats. It considers factors such as unauthorized access, data exfiltration, malware attacks, social engineering, and other malicious activities that can compromise the confidentiality, integrity, and availability of organizational data.

**Mitigation:** The d-TM model considers determining appropriate control as a part of the threat analysis process. Where the controls are determined to mitigate the identified threats, these controls represent a set of policies, procedures, techniques, or technology designed to reduce or eliminate the impact of cyber threats on the organization. Moreover, the model incorporates cybersecurity assurance as a key objective, driving the analysis and evaluation of determined threat controls.

**Assurance:** In the context of the analysis of the d-TM threat, the term "Assurance" denotes the extent of certainty regarding the implementation of security controls within an organization with the purpose of effectively minimizing potential threats. This allows organizations to acquire knowledge regarding the strengths and weaknesses of said security controls and make well-informed decisions aimed at enhancing their cybersecurity posture. The notion of assurance plays a significant role in this regard. Within the d-TM threat mitigation approach, assurance is provided by means of evaluation criteria for determining controls using various factors, distinguishing it from other approaches that solely aim to provide a list of potential controls for mitigating threats, without guaranteeing the effectiveness of said controls in mitigating the given threat. Nonetheless, d-TM undertakes an evaluation of the suitability of the suggested mitigation control, which evaluates the efficacy of security controls by taking into consideration three evaluation criteria: Completeness, Effectiveness, and Complexity. Control assurance instils confidence in an organization's cybersecurity mitigation strategy.

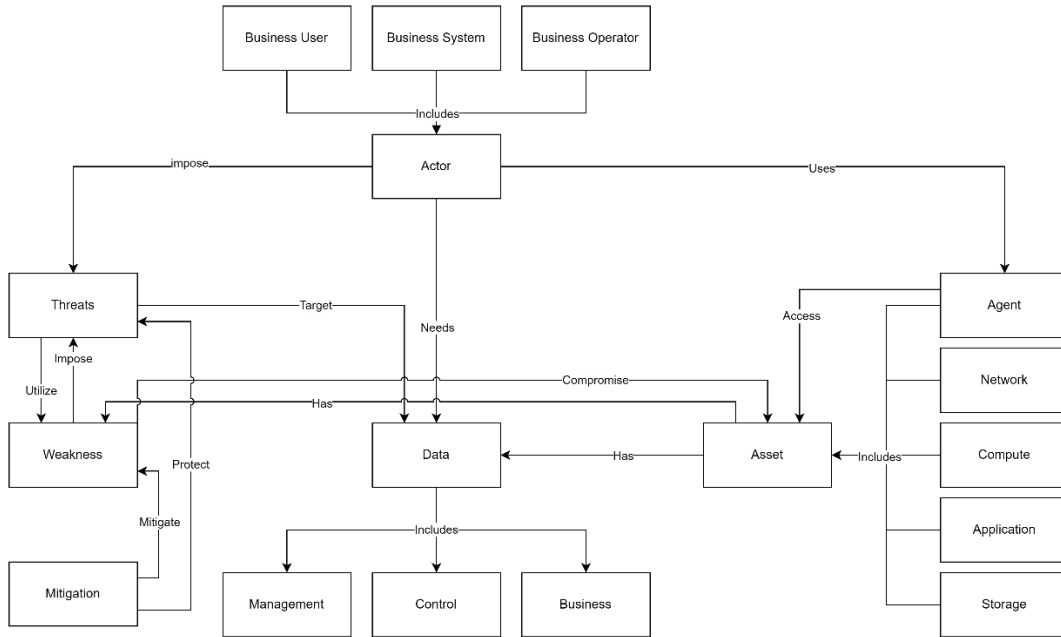


Figure 4.5. d-TM conceptual model

For instance, An organization provide a remote technical support services, relying on helpdesk platform as a business service to file a technical support cases. In such scenario as shown in figure 4.6, the conceptual model components can be presented and initiated as an Actor need to access helpdesk portal, using an agent which is a web browser for accessing the helpdesk portal. The agent exchanges data over networking assets i.e., switches that facilitate access to an application that is hosted on compute(server). The data could be stored in locally attached storage or remotely in network storage. Each of these assets could store, process or exchange data at any time, while data is categorized to three distinct types, specifically management data, control, and business. on other hand, the threat assessment diligently examines weaknesses within each data asset, as these weaknesses have the potential to jeopardize the security of data at various levels and stages. To bolster security assurance, it is imperative not only to identify these threats but also to take

proactive measures to mitigate them effectively. By comprehensively understanding the landscape of data vulnerabilities and strategically implementing security controls, organizations can fortify their defences and enhance the overall assurance of cybersecurity for their business operations.

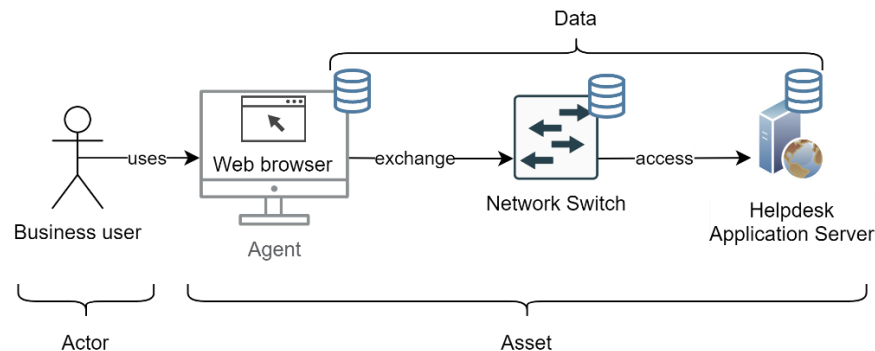


Figure 4.6. Example showing d-TM model components

## 4.5 d-TM Process

The process methodology used for threat analysis in the d-TM model comprises four main phases: data collection, data analysis, threat analysis, and threat mitigation. These phases are designed to leverage d-TM concepts and are critical in identifying, assessing, and managing threats effectively. **Data Collection** In the initial phase, relevant data is gathered to gain a comprehensive understanding of business processes, services, and assets. This data serves as the foundation for the threat analysis process of the d-TM model. **Data Analysis**, once data assets and services information are collected, this phase involves a thorough analysis of the data assets to determine data levels and phases that are used as a foundation for the next stages. Also, at this stage, the DFD of organizational infrastructure is constructed. The DFDs help identify the interconnections and relationships between infrastructure elements, providing insights into potential vulnerabilities and data access points. **Threat Analysis**: In this crucial phase, the identified assets and data are subjected to a comprehensive evaluation for weaknesses and related threats. The threat analysis process includes examining each asset and its associated data for potential weaknesses, which are then matched with known threats from the d-TM threat database. The identified threats are then prioritized based on their potential impact on business data. **Threat Mitigation** is the final phase that focuses on developing appropriate controls to address and reduce the prioritized threats. These controls are carefully selected based on their effectiveness in mitigating specific threats. The controls are also evaluated to ensure cybersecurity assurance for the organization. Figure 4.7 provides an overview of the d-TM approach, outlining the requirements, stages, and outcomes.



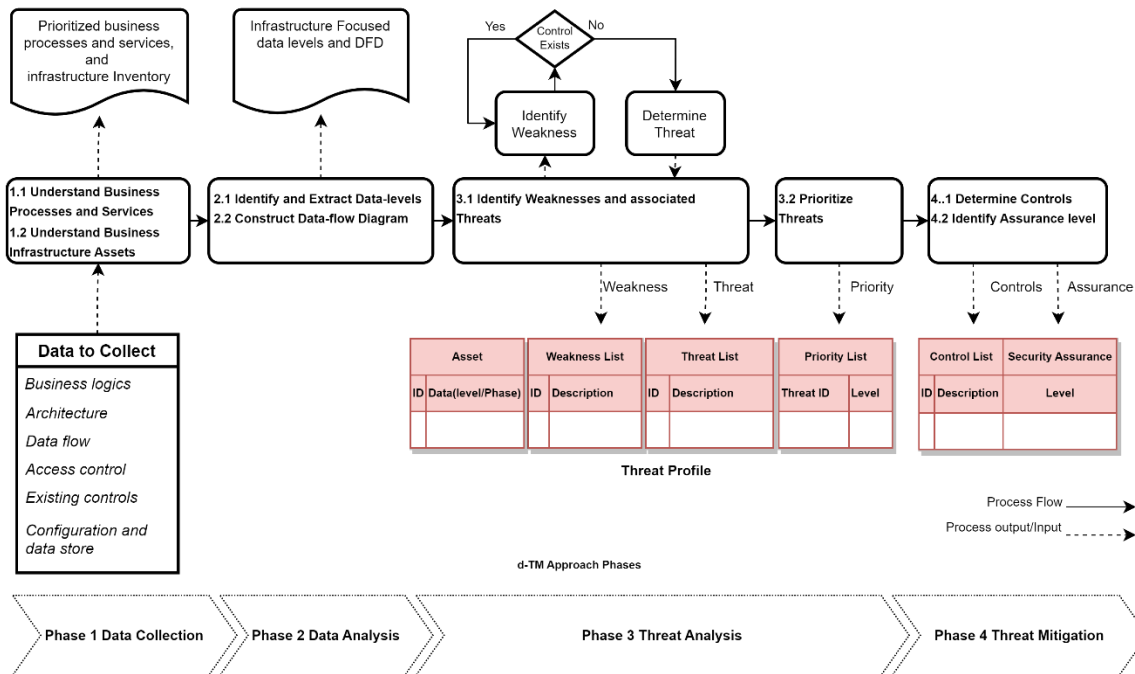


Figure 4.7. An overview diagram of the Data-driven threats analysis approach.

The sequential nature of these phases ensures a methodical and well-structured approach to threat analysis. The insights gained from each phase contribute to the development of a comprehensive threat profile table, consolidating all the information obtained throughout the analysis. By following this systematic methodology, organizations can proactively identify and address potential threats, strengthening their cybersecurity posture and safeguarding their valuable data assets.

The d-TM model is conducted using multiple activities, and each activity consists of a number of steps. Table 4.2 illustrated the d-TM four main activities, where each activity consists of two steps to achieve the intended threat activity process. The table also describes each activity pre-requests and the technique incorporated to extract the results. Furthermore, the table defines the intended stakeholder's role in participating in this activity, as well as the expected outcome of this activity. However, Within the scope of the research objectives, the four distinct activities constitute the integral components of the comprehensive automation process for threat analysis. Each activity, along with its subsidiary processes, is subjected to automation, with the intricacies of this procedure meticulously delineated in the forthcoming Chapter 5. This structured approach ensures a seamless and efficient progression towards a fully automated threat analysis system.

<b>Activity</b>	<b>Steps</b>	<b>Pre-Requisites</b>	<b>Technique/Method</b>	<b>Role</b>	<b>Output</b>
<b>Activity 1 Data collection</b>	1.1 Understand Business Processes and Services	Organization business stakeholders' contact information, roles, and responsibilities	interviewing business stakeholders	Organization's Business Stakeholders	List of high-priority business processes and related service functionalities
	1.2 Understand Business Infrastructure Asset	Organization technical stakeholders' contact information, roles, and responsibilities	Interviewing technical stakeholders and collecting asset inventory, designs, configurations, logs	Organization's technical stakeholders	Prioritized service Infrastructure components and operational data
<b>Activity 2 Data Analysis</b>	2.1 Identify and Extract Data-levels and phases	Infrastructure assets and operational data	Extracting asset operational data relevant to d-TM three data-levels and phases	Security Analyst	Relevant asset operational data mapped to d-TM data-levels and phases
	2.2 Construct a Data-flow Diagram	Details of services and infrastructure components	Use the DFD model to represent business services along with supporting infrastructure	Organization's technical stakeholders and Security Analyst	Presentation of business services in reference to d-TM using DFD format
<b>Activity 3 Threats Analysis</b>	3.1 Identify Weaknesses and Associated Threats	Assets operational data supporting services and threats dictionary	Analyse operational data with the aid of CWE and CAPEC catalogue	Security Analyst	A list of potential weaknesses and threats targeting each d-TM data-level
	3.2 Prioritize Threats	List of identified threats and relevant weaknesses	Use of d-TM scoring system to rank threats	Security Analyst	Prioritized list of threats
<b>Activity 4 Threat Mitigation</b>	4.1 Determine Controls	Identified threats and weaknesses	Use of the NIST catalogue to find suitable controls	Security Analyst	list of applicable controls to each threat
	4.2 Determine Assurance-level	Identified controls	Use of d-TM scoring system to rank control assurance	Security Analyst	Assurance levels of each identified controls

Table 4.2. Summary of d-TM Activities

## Activity 1. DATA COLLECTION

This phase collects data relating to the organizational digital services and business logic and develops an inventory of the organization's infrastructure assets and related data to provide a thorough understanding of critical business services and supporting infrastructure assets. It consists of two steps.

### Step 1.1. Understand Business Processes and Services

In this initial step, the focus is on identifying the business context, which encompasses the organization's main objectives, such as retail business, financial services, healthcare, etc. Underlying activities, known as Business Processes, are then recognized, and these processes aim to help the organization achieve its business objectives. Examples of business processes include sales, purchasing, marketing, and logistics.

Supporting these processes are digital services, referred to as Business Services, which enable the smooth execution of business processes. These digital services may include sales systems, marketing systems, HR systems, and more. Additionally, there is Business Infrastructure, which represents the cyber assets responsible for running the digital services, as illustrated in Figure 4.8. Examples of business infrastructure assets are servers and switches.

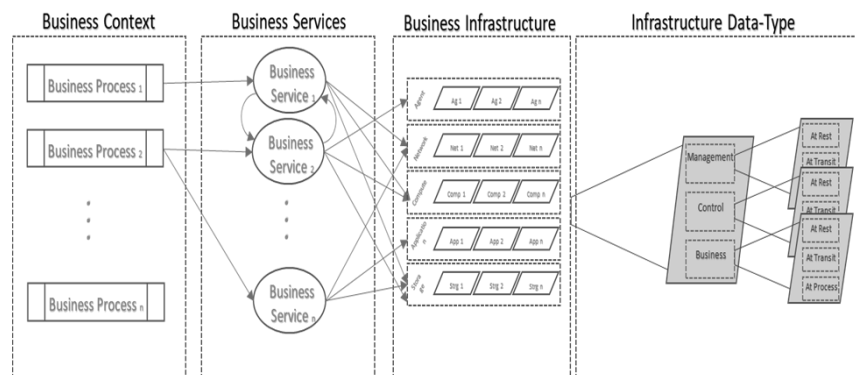


Figure 4.8. Business, services and infrastructure relationship mapping

Security analysts need to have a clear understanding of the business context and its underlying logic. To achieve this, collaboration and participation of various organizational stakeholders are crucial. This involves engaging with business decision-makers, systems architects, software developers, security administrators, and others.

The process of understanding the business context, processes, and supporting services involves conducting two levels of interviews: the strategic level and the operational/technical level. *Strategic-level* interviews target business decision-makers to identify business context, processes, and their priorities towards business continuity planning. On the other hand, *technical-level* interviews target the operational team to gain insights into business operational processes, digital services, and infrastructure.

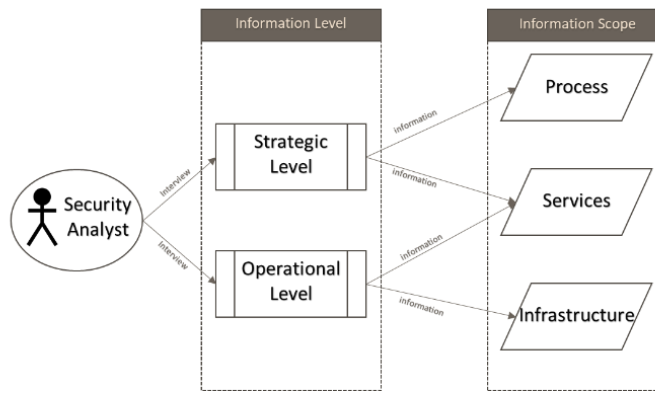


Figure 4.9. Business interview levels

The ultimate goal of this step is to define the business logic, outlining essential business processes, digital services, and the relevant infrastructure. Each functional process in the organization is supported by one or more digital services and corresponding infrastructure. The criticality of each selected service is assessed during the security analysis. The categorization of identified services in the threat analysis process is as follows:

- High (H): Represents mission-critical services that offer core functions to the business. The organization cannot function without these services.
- Medium (M): Represents supporting services to core functions where business interruption is limited and short-lasting.
- Low (L): Represents services that provide support to fundamental functions and allow the organization to operate with minimal effect for an extended period.
- Non-critical (NC): Represents services that support the business but do not directly influence core functions, such as a customer satisfaction system.

By carefully understanding and categorizing these business processes and services, the threat analysis process can focus on the most critical areas, enabling effective security measures to safeguard the organization's valuable assets and data.

### **Step 1.2. Understand Business Infrastructure Assets**

During this critical phase, the focus is on identifying and comprehending the infrastructure assets that play a pivotal role in supporting each critical service. The assets are carefully classified based on their type, management techniques, and inherent dependencies. To achieve this, the process actively involves technical experts by conducting interviews and gathering critical design materials. These interviews provide valuable insights into the technical aspects of the process, providing a deeper understanding of the system's architecture, requirements, and potential gaps. This holistic approach involves obtaining crucial information, such as system architecture, configuration files, logs, and other pertinent details.

The primary objective of this step is to empower security analysts to identify and assess the infrastructure components that are instrumental in supporting prioritized business functions. For this purpose, three properties are particularly considered for asset valuation:

- **Asset Type:** This property involves gathering essential information to characterize each asset, such as its name, model, and other relevant details. Additionally, to facilitate a comprehensive understanding of the assets, each of them is mapped to one of the five d-TM layers, which are as follows:
  - Agent (Agt)
  - Network (Net)
  - Compute (Cmp)
  - Application (App)
  - Storage (Stg)

Table 4.3 provides a detailed overview of the information provided by this property, streamlining the categorization and comprehension of the various infrastructure assets within the organization.

Asset	Description
Infrastructure Type ID	It represents the d-TM infrastructure types, including Agent, network, compute, application, and storage; also, it represents the identification number of the asset based on d-TM, for instance, such as Net0. It refers to the first identified network asset
Name	It represents the asset name in the organization infrastructure
Software/Hardware Model	It represents the software or hardware model number or brand
Software Version	It represents the software version number

Table 4.3. Asset type table

- **Asset Administration:** This property examines the access mechanisms used by administrators to operate the assets effectively. Security analysts are required to determine the protocols, tools, access conditions, and necessary privileges for operating each asset. Table 4.4 presents a comprehensive representation of the information provided by this property, aiding in the understanding of the administrative aspects of the assets.

Asset	Description
Mgmt. Protocols	It represents the management protocols/ports used to operate assets i.e., SH, RDP, HTTP/S, etc.
Mgmt. Tool	It represents the tool, system, or software used to operate the asset.
Mgmt. Access	It represents the access type required to operate the assets; there are three access types, direct access(physical), remote network access (internet/adjacent network i.e., WAN), and local network access (LAN).
Mgmt. Privilege	It represents the required user privilege to operate the asset; there are four access privileges, as below: <ul style="list-style-type: none"> <li>• User: domain user or local user</li> <li>• Admin: domain admin or local admin</li> <li>• Service: Local service or remote</li> </ul>

Table 4.4. Asset administration table

- **Asset Dependency:** Asset Dependency: Understanding the inherent dependencies among assets is of utmost importance in ensuring seamless service delivery. This property plays a

crucial role in identifying potential attack vectors resulting from weaknesses in dependent systems, leading to sequential failures or cascading breaches. For assessing asset dependency, d-TM adopts the MITIGATE Classification(*European Commission MITIGATE Project, 2020*), which considers two key factors: Dependency Type and Access. The Dependency Type defines the relationship between two assets, which could be categorized as host, exchange, storage, control, or process dependencies. On the other hand, the Access mechanism delineates how an asset communicates with another, with three types of access: direct, local, and remote. Table 4.5 provides valuable insights into the information gleaned from this property, enabling a comprehensive understanding of asset interdependencies.

<b>Dependency Type</b>	<b>Description</b>
Host	An asset is hosted by another asset, such as a virtual machine is hosted by a physical machine.
Exchange	An asset uses another asset to exchange information.
Storage	An asset uses another asset to store and retrieve data.
Control	An asset is controlled or managed by another asset.
Process	An asset uses another asset to utilize its processing capabilities or functionality.
<b>Dependency Access</b>	<b>Description</b>
Direct Access	It means the source is communicating with the destination using direct access or physical access.
Local Network	It means the source is communicating with the destination using the same local area network (LAN) or broadcast domain.
Remote Network	It means the source is communicating with the destination using a wide area network (WAN) or internet.

Table 4.5. Asset dependency table

## **Activity 2. DATA ANALYSIS**

In the previous phase, critical business services and supporting infrastructure are identified, and business logic is realized. This phase consists of two steps: discovering relevant data information while taking data levels and phases into account. The second step is to present the identified service assets together with a data flow diagram (DFD).

### **Step 2.1. Identify and Extract Data-levels**

This step builds upon the information gathered in the preceding phase, focusing on how data is processed at each asset for the identified important services. Security analysts with the necessary expertise examine technological configurations, codes, and designs that interact with data.

When extracting information, the d-TM model comes into play, guiding analysts to consider critical factors such as actors involved, specific layers within the d-TM framework, and the data levels being processed. The relevant details are then carefully recorded into a table, which may include configuration or code lines that describe specific functions. For assets with graphical interface-style configurations, such as Windows-based compute and some user applications, a

descriptive language alternative might be employed, providing insights to various stakeholders, including organization technicians, developers, and non-expert managers, about how assets are configured to process, transmit, or store data securely.

The exercise of data-level identification and extraction is security-driven, with a keen focus on aspects related to data processing, transmission, and storage. For instance, consider a computing device running a Linux operating system, accessed through the SSH protocol for management purposes. In this scenario, it becomes evident that the data being handled is management-level, with the actor being the business operator. The pertinent information or configuration details to extract might involve the SSH authentication technique, authorization level, protocol version, access port, and other relevant characteristics. Moreover, Modern applications often rely on API calls for seamless interactions. Extracting Application API configurations could be crucial for ‘control data’, particularly when business systems communicate with each other to exchange data or attributes. Table 4.6 serves as a valuable reference for the organization's security team, mapping out the precise points within the infrastructure where business data operates and the corresponding data phases.

Business Service (Bs)	Actor Use-case	d-TM Layers	Management Data-level (mD)	Control Data-level (cD)	Business Data-level (bD)	At Rest Data-phase (Dr)	In Process Data-phase (Dp)	In Transit Data-phase (Dt)
Bs <sub>x</sub>	Business-user (USR)	Agt <sub>x</sub>			D	●	●	●
		Net <sub>x</sub>			D			●
		Cmp <sub>x</sub>			D			●
		App <sub>x</sub>			D	○ <sup>1</sup>	●	●
		Stg <sub>x</sub>			D	○ <sup>1</sup>	○ <sup>1</sup>	○ <sup>1</sup>
Bs <sub>x</sub>	Business-Operator (OPR)	Agt <sub>x</sub>	D			●	●	●
		Net <sub>x</sub>	D			○ <sup>1</sup>	○ <sup>2</sup>	○ <sup>2</sup>
		Cmp <sub>x</sub>	D			○ <sup>1</sup>	○ <sup>2</sup>	○ <sup>2</sup>
		App <sub>x</sub>	D			○ <sup>1</sup>	○ <sup>2</sup>	○ <sup>2</sup>
		Stg <sub>x</sub>	D			○ <sup>1</sup>	○ <sup>1,2</sup>	○ <sup>1,2</sup>
Bs <sub>x</sub>	Business-System (SYS)	Agt <sub>x</sub>		D		●	●	●
		Net <sub>x</sub>		D				●
		Cmp <sub>x</sub>		D				●
		App <sub>x</sub>		D		○ <sup>1</sup>	●	●
		Stg <sub>x</sub>		D		○ <sup>1</sup>	○ <sup>1</sup>	○ <sup>1</sup>

Table 4.6. A table represents the data level and phase at any asset

Where “x” refers to service or asset\_id. “D” refers to the presence of the data level. “●” refers to the presence of the data phase. “○” refers to a potential presence of data; the presence depends on other factors. “1” The Data is stored locally (locally attached disk) or remotely, while remotely it needs data to be sent, processed, and stored at network storage. “2” The asset which the operator intends to access.

## Step 2.2. Construct Data-flow Diagram

In this step, we focus on creating enhanced Data Flow Diagrams (DFDs) for the prioritized services identified in the previous phase. DFDs are a widely employed technique in threat modelling as they offer a clear and simple graphical representation of how data flows through systems. However, in the d-TM approach, we aim to go beyond the traditional DFDs by incorporating data levels, actors, and the layered technique to make the diagrams more useful for comprehending data, not just flow.

The process begins by selecting the prioritized services from the previous phase. These are the critical business services with high or very high threats that demand immediate attention. Furthermore, each prioritized service is deconstructed into its functional layers, such as agent, network, compute, application, and storage. This layer-wise breakdown allows for a more detailed understanding of data interactions at distinct levels. While step 1 identifies each data level and phase, security analysts can refer to Table 4.6 to identify the specific data locations at each functional layer and phase of the service. However, this step provides insights into where data is stored and how it moves within the system.

Enhance DFDs with d-TM principles incorporate the symbols from Table 4.7 into the DFDs to represent components, data flows, data stores, actors, and processes. Additionally, leverage d-TM principles such as data levels and actors to enhance the diagrams further.


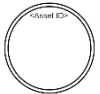



Item	Description	Symbol
Actor	Represent any actor demanding data.	
Infrastructure Asset	Represent asset object for particular asset id.	
Data at Rest	Represent asset object for stored data.	
Data in Process	Represent data in process for any asset object.	
Data in Transit	Represent data in transit for any asset object.	

Table 4.7. A d-TM Data flow diagram utilized symbols.

### Running Example of Deploying d-TM Enhanced Data Flow Diagram (DFD)

To illustrate the deployment of d-TM enabled DFD using a concrete example, Figure 4.10 highlights a DFD with d-TM enabled, providing a comprehensive visualization of data flow across different functional layers and phases for a prioritized service. In this example, we have a business use case scenario where a business user (B.user) accesses a critical business application.



- **User Interaction:** The business user (B.user) interacts with the system by accessing the business application. To access the service, the user leverages an Agent, represented by a web browser (Agt1) such as Mozilla Firefox.
- **Data Transmission:** The user's request is then transmitted to the organization's gateway router (Net2). The router routes the request to a specific compute (Cmp3), which is responsible for hosting the application (App4).
- **Data in Transit:** During the data flow, the information is considered in a transit state from the Agent to the application. This phase represents data being transmitted between the user's web browser and the application hosted on the computer.
- **Data Processing and Storage:** Depending on the required service, data may undergo processing and is stored either temporarily or permanently. The Agent and the Application are the components responsible for processing the data. Both of these components can store data temporarily during their operations.
- **Data Storage:** In this example, the application is linked to network storage (stg5), where data is stored persistently for future access or analysis. This storage represents the final destination for important data generated or utilized by the application.
- **Data Applicability and Asset Identification:** The DFD highlights essential information, specifically the data that is applicable to be saved in three potential assets: Agent, Application, and Storage. These assets represent different components of the system that handle, process, and store the data.

In summary, this d-TM-enabled DFD offers a clear and informative overview of how data flows through the different functional layers and phases of the prioritized service. It illustrates the interactions between the business user, the web browser (Agent), the application (hosted on the compute), and the network storage, with a focus on data transmission, processing, and storage points. By identifying the data applicable to each asset and understanding the data flow, organizations can assess potential vulnerabilities and prioritize security measures to mitigate threats effectively.

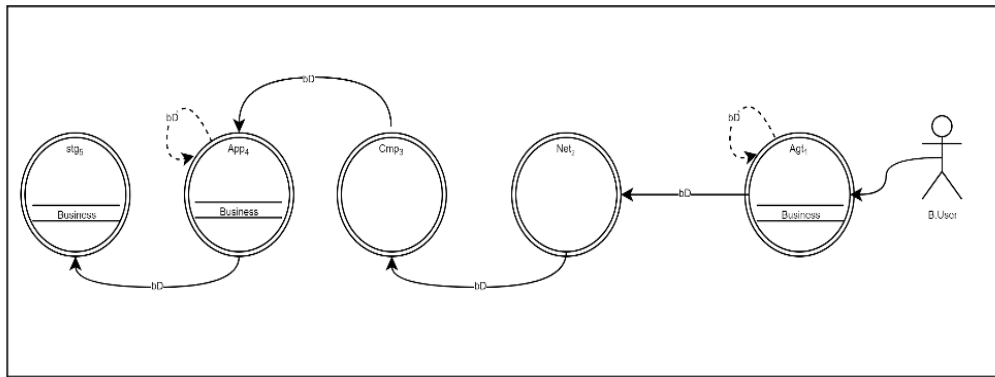


Figure 4.10. An example of a d-TM-enabled data flow diagram.

By following this meticulous data-level identification and extraction process, security analysts gain comprehensive insights into the data landscape of the organization. This understanding forms the foundation for effective threat analysis and allows for the creation of data flow diagrams (DFD) that illustrate the intricate paths and interactions of data throughout the system. Through the combination of d-TM principles and data analysis, potential vulnerabilities can be identified and addressed, ultimately strengthening the organization's data protection and resilience against cyber threats.

### ***Activity 3. THREATS ANALYSIS***

Activity 3 encompasses crucial threat assessment activities, comprising two steps: identifying weaknesses and associated threats and prioritizing the detected threats. The outcome of this phase is documented as a threat profile knowledgebase, providing organizations with a comprehensive overview of threat characteristics and weaknesses that could be exploited to compromise data assets.

#### ***Step 3.1. Identify Weaknesses and Associated Threats***

In this initial step, the focus is on investigating the collected data at each data level to identify potential weaknesses that may lead to threats. Security analysts employ methods such as manual code review and architecture or design review to detect weaknesses in the acquired information. While manual approaches are effective, they might be resource-intensive. In many cases, using a technique like architecture or design review can offer cost-effective and comprehensive coverage for networked assets. Expert judgment is vital in this process to identify suitable weakness identifiers, considering that it is impractical to be familiar with all technologies.

The examination aims to find potential weakness identifiers by comparing the acquired data with possible scenarios where assets could abuse or compromise data at each data level. The d-TM model relies on the Common Weakness Enumeration (CWE) knowledgebase by MITRE as a reference for weaknesses. CWE provides multiple views that aid expert analysts in finding matching weaknesses based on software, hardware, research concepts, and other criteria. Once weaknesses are identified, they are recorded with respective CWE IDs for each data level and phase. Once weaknesses are detected, the next step is to identify threats associated with each weakness. For this, certain conditions are used to determine the existence of a threat. If a threat exists for any use case, each weakness is evaluated to discover the associated threat. Similar to the weakness identification process, the d-TM model encourages subject matter experts to adapt identified threats into a common security language. The research relies on the MITRE knowledgebase's Common Attack Pattern Enumeration and Classification (CAPEC) as a reference for documenting threats. The correlation between CWE IDs and CAPEC IDs simplifies the process of matching threats to identified weaknesses, streamlining threat assessment. With all potential threats and weaknesses linked to matching CWE and CAPEC IDs, the outcome is documented as a result of the threat profile knowledgebase. The knowledge base serves as a valuable resource for the organization, providing a clear understanding of the threats and weaknesses that need to be addressed to enhance data asset security.

## Running Example of Identifying Weaknesses and Associated Threats

In the visual example of applying d-TM steps to a business service, Figure 4.11 represents a helpdesk system chosen as a high-priority service for the business. The service is divided into d-TM proposed layers, each mapped to a specific asset identifier. Let us analyse the findings and weaknesses in the network layer (Net0) based on d-TM data levels:

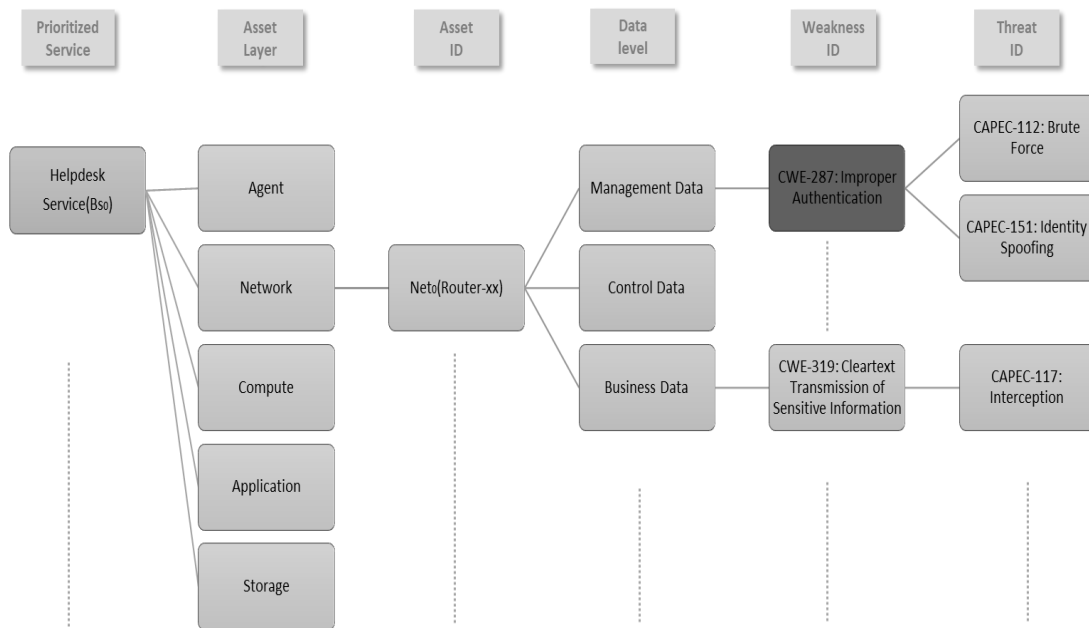


Figure 4.11. A visual presentation of d-TM weakness and threat analysis

The analysis summary is as follows:

### 1. Asset Identification:

- Asset: Net0 (Router)
- Layer: Network Layer

### 2. Weakness Identification:

- Weakness: CWE-287 ("Improper Authentication")
- Description: The assessment identifies a weakness related to improper authentication in the management data within the network layer. This indicates a vulnerability in the way authentication mechanisms are implemented in the router.
- Weakness: CWE-319 ("Cleartext Transmission of Sensitive Information")
- Description: Another weakness is found in the handling of business data within the network layer. Specifically, it relates to the transmission of sensitive information in cleartext, which means that data is not encrypted

during transmission, making it susceptible to interception and unauthorized access.

### 3. Threat Identification:

- Threat: Brute Force (CAPEC-112)
- Description: The first weakness identified (CWE-287) in the management data is associated with the threat of a Brute Force attack (CAPEC-112). This threat involves an attacker attempting to gain unauthorized access to the router by trying various username and password combinations until they find the correct one.
- Threat: Identity Spoofing (CAPEC-151)
- Description: The same weakness (CWE-287) is also associated with the threat of Identity Spoofing (CAPEC-151). This threat implies an attacker impersonating an authorized user or device to gain unauthorized access to the router.
- Threat: Data Interception (CAPEC-117)
- Description: The weakness in handling business data (CWE-319) makes the business data susceptible to Data Interception (CAPEC-117). This threat involves unauthorized individuals intercepting and capturing sensitive information as it is transmitted over the network.

In summary, the assessment of the network layer (Net0) in the helpdesk system using d-TM identifies two weaknesses: "Improper Authentication" (CWE-287) and "Cleartext Transmission of Sensitive Information" (CWE-319). These weaknesses pose potential threats to the management data and business data within the network layer. The threats associated with the weaknesses are Brute Force (CAPEC-112), Identity Spoofing (CAPEC-151), and Data Interception (CAPEC-117). Understanding these weaknesses and threats is crucial for developing effective security measures to protect the network layer and the overall helpdesk system from potential attacks and compromises.).

#### **Step 3.2. Prioritize Threats**

This step aims to determine the criticality of each threat to the business, enabling the development of an effective control strategy to address threats appropriately. The prioritization of threats involves a novel approach that considers multi-factor evaluation, including Business-as-Target (Bt), Threat-Complexity (Tc), and Business-Impact (Bi). The overall priority is determined based on the correlation of these three factors across five threat scales: Very High (VH), High (H), Medium (M), Low (L), and Very Low (VL).

- **Very High:** It represents an extremely high impact on business continuity and critical business services which require immediate action.
- **High:** It represents a significantly high threat to business continuity and critical business services due to its impact and need to act in a brief time frame

- **Medium:** It represents an intermediate threat to business continuity, and no critical business services get interrupted due to its impact. Furthermore, some supporting business services are impacted, and business can run for a longer time. Also, application action should be within a year’s time frame.
- **Low:** It represents a low threat to business continuity, and no critical business services get interrupted due to its impact on options control. Furthermore, some supporting business services are impacted, and business can run normally.
- **Very Low:** It represents a significantly extremely low threat to business continuity, and no critical business services and supporting services get interrupted due to its impact. Also, applying for action could be optional or ignored.

The three factors used in the prioritization process are described below:

**Factor 1. Business-as-Target (Bt)**

This factor assesses the likelihood of an organization being highly targeted by an attack. It considers the correlation between the threat-occurrence metric and the attacker-gain metric. The threat-occurrence metric represents the likelihood of a specific threat occurring based on the organization's attack history and public threat awareness records, e.g., the IBM X-Force Threat Intelligence Index(*IBM Security X-Force Threat Intelligence Index 2023, 2023*). While the attacker-gain metric reflects the gain value behind the attack.

- **Metric 1. Threat-Occurrence** represents the probability of a particular threat occurring for an asset in three scales:
  - High(H): The organization experiences this threat twice or more in a one-year timeframe. Industry researchers forecast this threat as a top-rated attack for similar businesses within one year.
  - Medium(M): The organization experiences this threat once in a one-year timeframe. Industry researchers forecast this threat as Medium-rated attacks for similar businesses within one year.
  - Low(L): The organization experiences this threat once or none in a two-year timeframe. Industry researchers forecast this threat as a medium to low-rated attack on similar businesses within the last two years.
- **Metric 2. Attacker-Gain** represents the goal behind the attack, including curiosity, personal gain, personal fame, or national interests, with three scales: high, medium, and low.

The two metrics correlation matrix for Factor (Bt) is presented in Table 4.8 as follows:

Bt	Threat-Occurrence likelihood		
	High	Medium	Low
Attacker-Gain Scale			
High	H	H	M

Medium	H	M	L
Low	M	L	L

Table 4.8. Business -as- target matrix

### Factor 2. Threat-Complexity (Tc)

This factor evaluates the complexity of a particular threat, considering Attacker-Capability and Access-Complexity. Attacker-Capability refers to the overall expertise, knowledge, and resources of the attacker to exploit a weakness, while Access-Complexity gauges the level of complexity involved in exploiting a weakness, considering the organization's access controls. The two metrics and correlation matrix are defined below:

- **Metric 1. Attacker-Capability** refers to overall attacker capabilities such as skills, knowledge, opportunities, and resources that the attacker incorporates to exploit a weakness. Capabilities could be estimated as High, Medium, and low.
  - High: A sophisticated level of expertise and knowledge with adequate resources for generating opportunities for continuous attacks
  - Medium: Moderate level of expertise and knowledge with reasonable resources for the considerable ability to generate multiple opportunities and continuous attacks
  - Low: A low level of expertise and knowledge with limited resources and the ability to attack
  
- **Metric 2. Access-Complexity** determines the level of complexity to exploit a particular weakness, where each organization has various levels of access and controls. Attackers often estimate the level of complexity of any attack to find the easiest and success-guaranteed approach to compromise organization data. Likewise, security analysts evaluate existing access and security mechanisms to understand the access complexity that could reduce the likelihood of exploiting existing weaknesses. Access complexity can be estimated as High, Medium, and low. Table 4.9 represents the correlation matrix for the likelihood of threat complexity to a particular asset.
  - Multi-level Access: The attacker requires a restricted access condition, and this condition requires an elevated level of effort and expertise that could go over a multi-stage of attack.
  - Single-level Access: The attacker requires a somewhat restricted access condition, and this condition requires a medium level of effort and expertise.
  - Direct: The attacker requires no restricted access condition.

Tc	Access-Complexity levels		
Attacker-Capability levels	Multi-level	Single-level	Direct
High	M	L	L

Medium	H	M	L
Low	H	H	M

Table 4.9. Threat complexity matrix

**Factor 3 Business-Impact (Bi)**

This final factor determines the impact of the threat on the business. It combines the results of Factor 1 (Bt) and Factor 2 (Tc) with the business impact probability, classifying impact as High (H), Medium (M), or Low (L).

- High(H): The expected impact of the mission-critical services that provide core functions to the business is High; the business cannot run without it.
- Medium(M): The expected impact on the business is medium, where supporting service is impacted; also, the business could run for some time.
- Low(L): Represents the expected impact to the business is low, where the business is run with minimal impact.

Based on the correlation of these factors, the overall threat priority is determined. Threats with a higher priority (e.g., VH or H) require immediate action to ensure business continuity and safeguard critical services. Lower-priority threats (e.g., M, L, or VL) might have less severe impacts, and the organization can plan actions accordingly. The overall correlation matrix of the three factors is illustrated in Table 4.10.

Threats Priority	Bi								
	High			Medium			Low		
Tc \ Bt	Low	Medium	High	Low	Medium	High	Low	Medium	High
High	VH	VH	H	H	H	M	M	M	L
Medium	VH	H	M	H	M	L	M	L	VL
Low	H	M	M	M	L	L	L	VL	VL

Table 4.10. A table represents the Matrix of threat priority

Finally, the complete threat analysis activity is summarized in a threat profile table, including prioritized threats, weaknesses, and relevant attributes. The threat profile provides the organization with a communication tool for the threat assessment process, allowing the organization to prioritize actions based on high-potential threats to the business. This structured and comprehensive approach enhances the organization's ability to address potential threats and weaknesses in its data assets proactively.

**Activity 4. Threat Mitigation**

This is the last phase of the proposed d-TM process, aimed at determining suitable controls to mitigate identified threats and ensure the organization's security objectives are achieved. This

phase involves interpreting and measuring security controls to assess cyber security assurance. d-TM considers multiple factors to ensure the appropriate level of assurance. The process consists of two steps:

#### **Step 4.1. *Determine Controls***

This step focuses on providing appropriate controls based on the identified threats for each data element. These controls are derived from weaknesses that attackers could exploit. d-TM advocates evaluating suggested mitigation guidelines by Common Weakness Enumeration (CWE) to identify controls that eliminate or reduce the threat's occurrence. Additionally, the Common Attack Pattern Enumeration and Classification (CAPEC) could suggest actions to mitigate the threats themselves. This step can be achieved using two actions as follow:

- **Understand Weakness Characteristics**

This step aims to gain a detailed understanding of the weakness causing the threat, including the required scope of action and the impact on data. CWE guidelines are used to obtain information about each weakness, particularly the "Scope" and "Impact" properties found under the "Common consequences" section. The "Scope" property categorizes weaknesses into domains such as confidentiality, availability, integrity, non-repudiation, and access control. The "Impact" property provides a brief description of the attacker's intended goal. This information provides insight into the scope of required mitigation. Additionally, security analysts need to identify the appropriate mitigation technique to address each weakness. The "Potential Mitigations" section in CWE offers guidance on potential techniques, serving as a guideline rather than a definitive solution. Security analysts analyse the provided information to develop the necessary understanding for selecting the most suitable control in the next step.

- **Analyse and Map Applicable Controls**

Building upon the scope understanding gained in the previous steps, security analysts enumerate each relevant NIST (National Institute of Standards and Technology) security control family. They consider the scope, impact, and mitigation techniques provided by CWE guidelines. Initially, potential candidate controls that align with the weakness's scope are identified. Then, the controls are assessed for their ability to protect against the weakness's impact. It is crucial to ensure that the security controls align with the mitigation technique guidelines provided by CWE. Once suitable security controls are determined, the mapping between weaknesses and controls is documented to the relevant threat identifiers.

#### **Step 4.2. *Determine Assurance-level***

The ultimate step in Activity 4 involves determining the assurance level of overall cyber security by considering the identified threats, controls, and data. The assurance level is based on the completeness, effectiveness, and complexity of the security controls in addressing threats to data security. These three factors are correlated to determine the overall assurance level of the control. Equation 1 is used to calculate the Overall Assurance Level (OAL), where OAL is considered High when the OAL value ranges from 7 to 9, Moderate when it ranges from 4 to 6, and Low when it is less than 4.



$$\text{Overall Assurance Level (OAL)} = \text{Completeness} + \text{Effectiveness} + \text{Complexity} \quad (\text{Equation 1})$$

- **Completeness (Ct):** Completeness refers to the extent to which the identified controls are comprehensive and relevant in addressing the identified threats and related weaknesses. During the evaluation process, controls are carefully assessed based on their ability to provide coverage that eliminates or minimizes the impact of specific threats. Table 4.11 categorizes controls into distinct levels of coverage: low, partial, or full coverage for specific threats. This evaluation helps provide valuable insights into the selection of suitable controls that effectively mitigate threats at specific data levels and phases.

A thorough analysis of completeness enables organizations to prioritize their control implementation efforts. By understanding which controls offer the most significant impact, security teams can efficiently allocate resources and focus on measures that align with their security objectives. It ensures that all relevant threat vectors are adequately addressed, reducing the organization's overall risk exposure, and enhancing cybersecurity resilience.

Completeness Description	Scale
The control provides the necessary features to mitigate the threat's likelihood without requiring any additional enhancements or supporting controls.	H (3)
The control provides some features to reduce the threat likelihood; however, additional enhancements or supporting controls are required.	M (2)
The control provides a significantly minimal feature to reduce the likelihood of the threat; however, additional enhancements or supporting controls are required.	L (1)

Table 4.11. Controls Completeness levels

- **Effectiveness (Ef):** Effectiveness evaluates the capability of security controls to perform their intended roles of protecting, detecting, and responding to threats. To be truly effective, a control must excel in all three aspects. Firstly, it should be capable of preventing threats from occurring in the first place, reducing the likelihood of successful attacks. Secondly, it must possess the ability to promptly recognize and detect any active threats within the system, enabling timely response and containment. Lastly, a well-performing control ensures a swift and efficient response mechanism, minimizing the impact of successful threats and facilitating recovery.

Table 4.12 provides guidance on evaluating the effectiveness of controls. By carefully considering this factor, organizations can ensure that their chosen controls are robust and capable of defending against a wide range of threats. Effective controls provide a strong defensive posture, reducing the risk of successful cyber-attacks and safeguarding critical assets and data.

Effectiveness Description	Scale
Control aims to prevent the occurrence of an attack, as well as to detect and respond when necessary, without the need for any further enhancements or supporting controls.	H (3)
Control aims to provide two essential roles, such as protecting and detecting the occurrence of an attack, with no response. However, further enhancements or supporting controls are required.	M (2)
Control aims to provide a single essential role, such as detecting the occurrence of an attack, with no protection and response. However, further enhancements or supporting controls are required.	L (1)

Table 4.12. Controls Effectiveness levels

- Complexity (Cx):** Complexity assesses the ease of implementing and operating a control within the organization's existing infrastructure. Controls that are overly complicated to integrate can pose challenges for the security team, potentially leading to errors or misconfigurations. Therefore, it is essential to select controls that can be smoothly integrated into the existing organizational systems without necessitating major changes. The evaluation of control complexity within the organizational context is indeed subjective and can vary significantly based on the specific expertise and infrastructure of an organization. In addressing this concern, it's important to recognize that this process is typically a collaborative effort involving multiple stakeholders within the organization. This includes the security team, IT department, and potentially external consultants who have a comprehensive understanding of the existing infrastructure and the skill set of the personnel involved. However, organizations may adopt a systematic evaluation process that includes criteria such as the required infrastructure changes, compatibility with existing systems, and the learning curve for the security team.

Furthermore, operation complexity becomes a concern when the security team lacks the necessary knowledge or expertise to handle the control effectively. In such cases, controls with clear and intuitive workflows are preferred, enabling the team to apply necessary actions efficiently. Table 4.13 offers insights into evaluating the complexity of controls. By opting for controls with manageable complexity, organizations can avoid potential implementation hurdles and ensure that their security measures remain practical and sustainable.

Complexity Description	Scale
Control can integrate seamlessly into organization infrastructure, and the team have the skills to implement and operate the control.	L (3)
Control can integrate into organization infrastructure with minimal changes, and the team have no skills to implement and operate the control.	M (2)
Control is complex to integrate into the organization infrastructure, and the team have no skills to implement and operate the control.	H (1)

Table 4.13. Controls Complexity levels

By considering these d-TM factors, organizations can make informed decisions when selecting and implementing controls, thereby enhancing their overall cybersecurity assurance. A comprehensive evaluation of completeness, effectiveness, and complexity ensures that controls align with the organization's security objectives and offer a robust Defence against evolving cyber threats. Eventually, d-TM provides a superior value to other threat models, including PASTA and STRIDE, by providing control assurance in the process of threat analysis.

- d-TM: d-TM's control evaluation strategy directly assesses the effectiveness of security controls by considering Completeness, Effectiveness, and Complexity. It offers a holistic view of an organization's cybersecurity assurance level.
- PASTA: PASTA concentrates on analysing attack scenarios and developing mitigation strategies but lacks a direct control evaluation component similar to d-TM's approach.
- STRIDE: STRIDE is a threat-based methodology that helps identify and address specific threats without explicitly evaluating control effectiveness like d-TM.

In summary, d-TM's control evaluation strategy is unique in its focus on assessing the overall cybersecurity assurance level by considering multiple factors related to control coverage, effectiveness, and complexity. This comprehensive approach enables organizations to gain insights into the strengths and weaknesses of their security controls and make informed decisions to enhance their cybersecurity posture. While PASTA and STRIDE are valuable threat modelling methodologies, they do not emphasize control evaluation in the same manner as d-TM.

## 4.6 Conclusion

This chapter represents one of the significant contributions to this thesis, which is introducing the d-TM approach along with its foundational pillars. At its core, this model places a strong emphasis on data, recognizing it as a pivotal component, and delves into the identification of weaknesses within data assets that are integral to an organization's data operations. By proactively identifying these weaknesses, the d-TM provides a strategic approach to uncover and address potential risks to the business proactively. Within the d-TM's threat analysis framework, three pillars stand prominently: data levels, phases, threat layers, actors, and a common security knowledge base. These pillars collectively equip the d-TM model with the necessary components to comprehensively evaluate the complexity of an organization's infrastructure while bridging this understanding with its core business functions. Moreover, the incorporation of a common security knowledge base serves as an invaluable resource, granting access to the latest industry insights regarding identified weaknesses, threats, and their corresponding mitigation techniques.

The construction of the d-TM model is supported by a systematic sequence of four activities, commencing with a thorough comprehension of the organization's business logic and concluding in the identification of critical threats that pose potential risks to business continuity. An integral feature of the d-TM is its threat mitigation capability, embedded within the threat analysis process, aimed at reducing or eliminating these identified risks. In essence, the d-TM model, along with its foundational components, aspires to provide essential insights into the threats that target

organizations and ensure the cybersecurity assurance of these businesses. This approach represents a noteworthy contribution to the realm of cybersecurity.

## CHAPTER *FIVE*: d-TM AUTOMATION

## 5.1 Introduction

Threat modelling is an essential tool for any organization that is serious about protecting its digital infrastructure. There are several threat modelling frameworks available, but they can be time-consuming and labour-intensive to implement. The data-driven threat modelling (d-TM) process is a more efficient and effective way to identify and mitigate threats. d-TM uses a data-driven approach to analyse and detect pre-attack activities, which can then be used to prevent future attacks.

This chapter introduces a d-TM tool that can help organizations automate the threat modelling process. The tool provides a workflow that guides users through the process of defining business processes, services, and IT infrastructure, conducting threat assessments, and developing mitigation strategies. The tool also generates threat profile reports that can be used to make informed decisions about security.

The d-TM tool is designed to help organizations:

- Identify critical services and data assets.
- Detect weaknesses and threats.
- Determine critical threats and proper controls.
- Assess the effectiveness of controls.

By using the d-TM tool, organizations can reduce the complexity and time consumption of manual threat analysis practice to automated process that reduces time and efforts for detecting the risk of cyberattacks and improve their overall security posture.

## 5.2 d-TM Platform Overview

The d-TM methodology has been implemented as the d-TM platform, a software application that helps organizations conduct security threat analysis, particularly on data. The platform provides a comprehensive workflow that guides users through the process of identifying business-critical services, identifying critical data assets, assessing weaknesses, identifying impactful threats, and investigating controls.

The d-TM platform is written in Python, Java, and HTML, and uses the Django web application module. This makes it compatible with any standard web browser and allows it to be used on a local network or the internet. The platform also uses a MySQL database to store collected and generated threat analysis data. The user interface of the platform is comprised of a number of web pages that present backend applications that are associated with the key d-TM processes, such as the data collection process. The platform's architecture is based on the use of multiple sub-applications, which provide adequate separation between the platform's many features and functions, as well as greater adaptability to modifications made to individual sub-applications.

## 5.3 d-TM Platform Design

The d-TM platform is developed in accordance with the standard guidelines for Django, which dictates that the main project should come first, followed by the assigned sub-applications. The

three sub-applications that support the platform are authentication, data collection, and data threat analysis. Each subsidiary application is connected to achieve the four d-TM activities. The authentication application is used to secure and manage user access to the platform, as well as the signup process and user profiles. The data collection application is the second in the process, where users feed the tool with required information about services and associated infrastructure. Lastly, the threat analysis application is where the core function of threat analysis is performed.

The overall design components of the d-TM platform, including applications and the primary web pages, are displayed in Figure 5.1

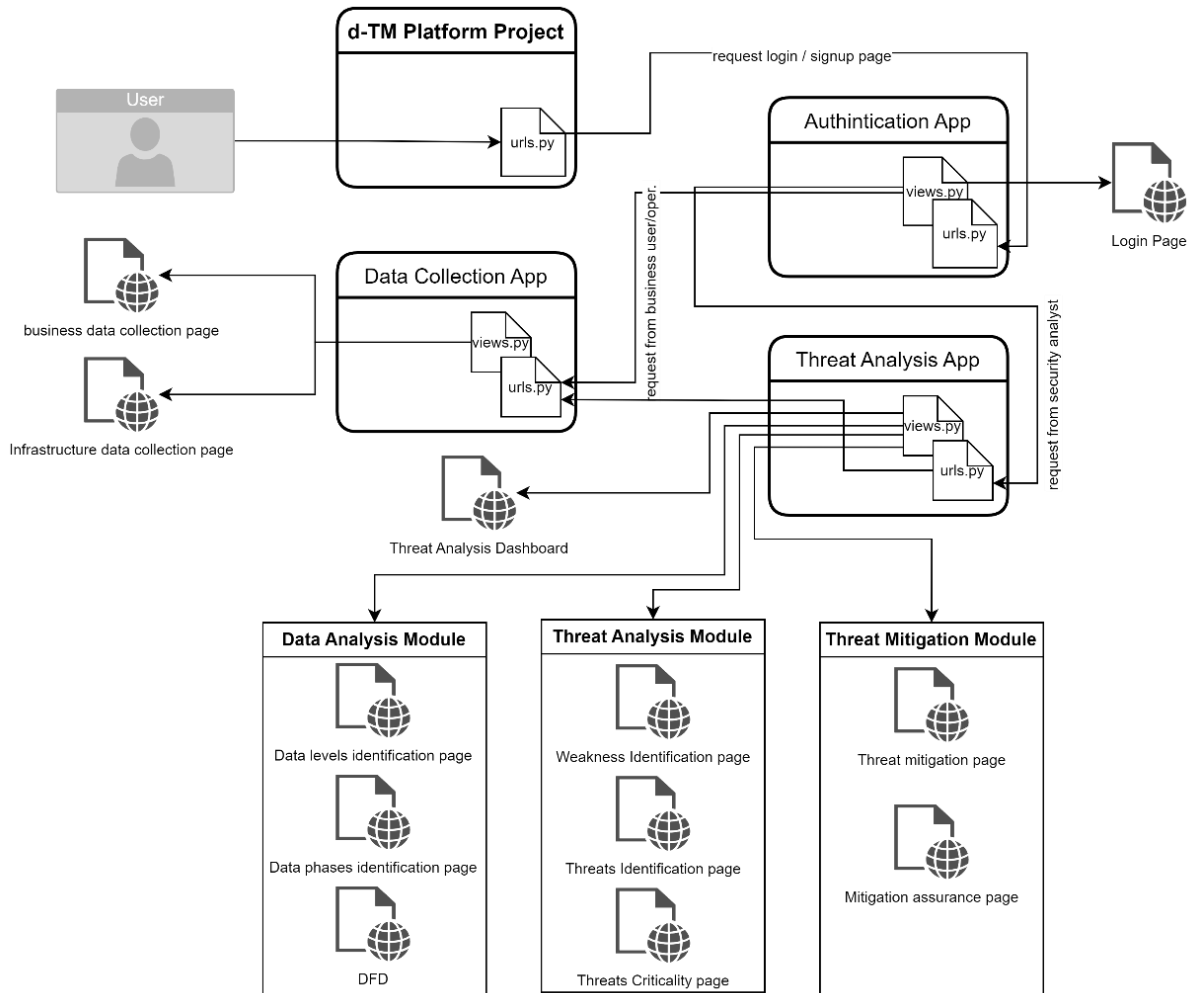


Figure 5.1 d-TM platform architecture

## 5.4 d-TM Platform Applications

The d-TM platform is constructed based on three applications, authentication, data collection, and threat analysis application. These applications are accessed sequentially, starting from the authentication process to access the platform till the threat analysis outcome is produced.

### 5.4.1 Authentication Application

The authentication feature is a vital component of the d-TM platform. It controls users' access to the platform's functionalities. Users can log in to the platform, and there is also a sign-up option for users accessing the platform for the first time.

d-TM was developed to accommodate multiple users simultaneously and classify users according to their organization and role. The application provides users with two distinct sign-up and sign-in forms:

- The sign-in form is for previously registered users and requires a username and password.
- The sign-up form is for collecting information about the user before they use the platform. To sign up, users must provide information such as their first and last name, organization, role, and unique user credentials (username and password).

The authentication functionality was developed securely, with the application being isolated from other modules in terms of its management and functionality. Additionally, a separate database stores users' credentials within the application. When a user accesses any application, they must first be authenticated and authorized before accessing the application. Since the authentication module is required for every application, every application must have one. Figure 5.2 shows a screenshot of the d-TM platform's login page, sign-up page, and user profile page. The user profile is located in the top right corner of all platform pages which provides information regarding the user that is referenced with provided data and enables the user to make any necessary updates to his information. The platform access link and credentials are detailed as below:

**URL:** <http://dtm-platform.ddns.net:8000/login/>

**Username:** dtm001

**Password:** Dtm@123456

**Important:**

- Please note that '**Identify Data level**' link in the Data Analysis sidebar conducts multiple comprehensive analysis and tuning for all imported configuration files, which is why the page to be loaded fully **it requires 30-45 minutes** (not a bug, it is due to browser constrains to number of loaded items However, this issue does not exist in Safari Browser).
- The tool shows the best visualization on Mozilla Firefox browser. However, other browsers work fine, but some visualizations could not load with proper view.



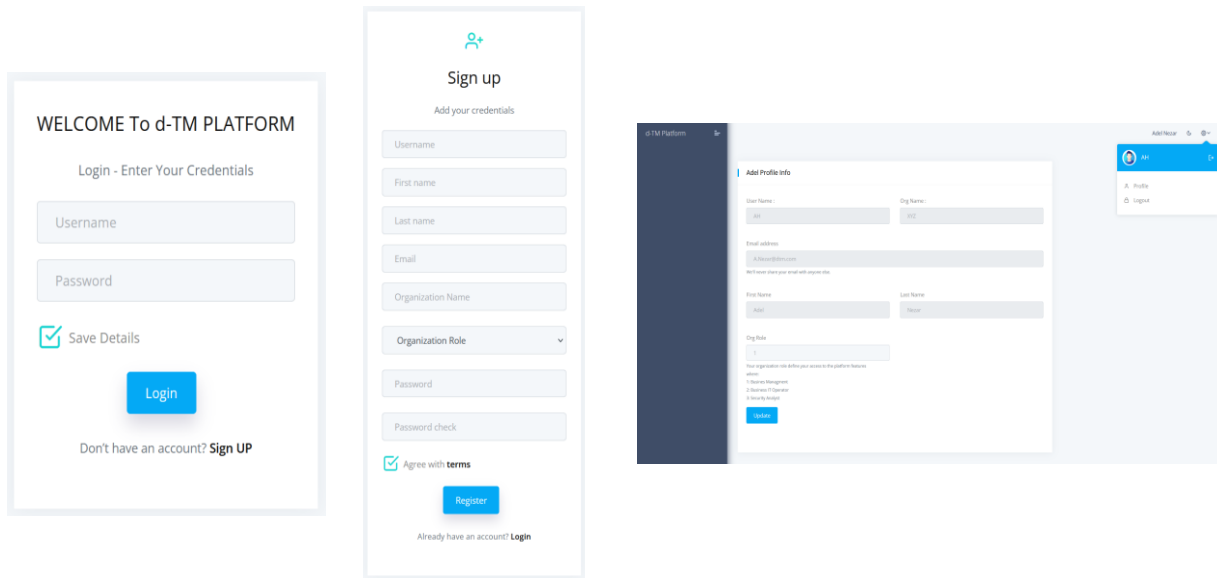


Figure 5.2 d-TM platform login, signup and profile page

### 5.4.2 Data Collection Application

The information required for d-TM threat analysis is obtained through the use of the data collection application. The application collects information about business services and evaluates their criticality using survey forms. Two distinct types of interface forms are provided for business stakeholders to use. The first form focuses on organizational digital assets, while the second concerns business services. However, the system can route the user's request to the appropriate data collection form; hence, Users must disclose their role within the company before using the system.

The first interface, shown in Figure 5.3, allows the business management team to register information about the running services of the business, including the criticality of those services to the continuity of the business. The user interacts with the system by filling out a single form that collects information such as the service name, description, and the criticality of the service. The form can be used to collect one or more services. However, the system will assign a unique identification number to each collected service. After the user has provided the service's specifics, they can submit the information to the system database. Furthermore, the web page provides the inventory list of collected services at the bottom of the page.

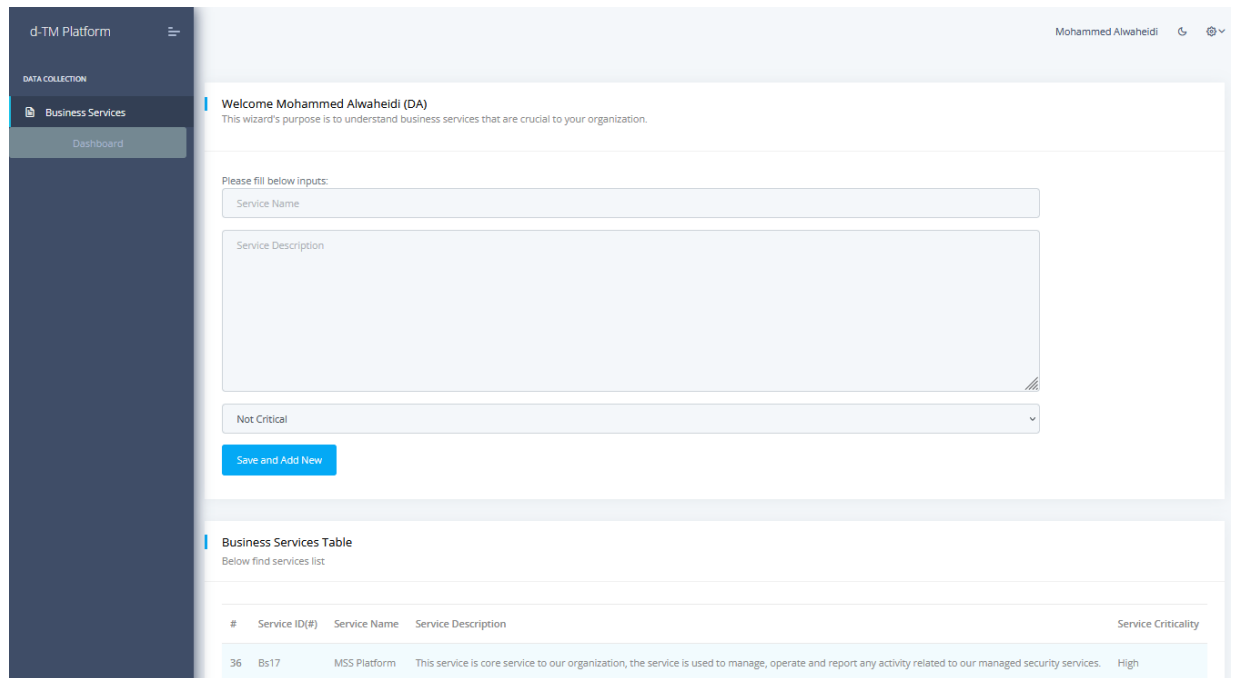


Figure 5.3 d-TM platform business data collection interface

Once the platform collects the services, the second step of the data collection is started. IT infrastructure administrators must sign up to the platform to provide details about infrastructure assets supporting each collected service. The process begins by selecting the business service for which this evaluation is intended and ends with an inventory table that displays all the collected assets. This process consists of three sequential input forms:

**Form 1: Data Asset Details**

This is the first form in the process of collecting data asset details. As illustrated in Figure 5.4, the form starts by selecting the service from the pull-down menu; then, the user must fill out all the form fields for each asset. These fields contain the asset name, asset ID according to the d-TM reference layers, and the ability to upload files, allowing users to share asset configuration files.

d-TM Platform Mohammed Alwaheidi

DATA COLLECTION

Business Infrastructure

Dashboard

---

**Welcome Mohammed Alwaheidi (DA)**  
 This wizard's purpose is to understand infrastructure supporting business services that are crucial to business services.

Please fill below inputs:

**Business Service**  
 Please select business service..

**Data Assets Details**  
 Please consider the asset to be provided in reference to real data flow sequence

Asset Type details

Agent

Asset Name	Asset Brand	SW / Ver
<input type="text"/>	<input type="text"/>	<input type="text"/>

Please upload configuration file:  
 No file selected.

**d-TM Layers** Please consider the below d-TM layers to provided real data flow sequence

d-TM Layers

---

**Service Data Assets Table**  
 Below find assets list

#	Service. ID(#)	Asset. ID(#)	Asset. Name	Asset. Brand	Asset. SW Version	Asset Config File
24	MSS Platform (Bs17)	Net1	Core-Switch-01	Juniper	21.2	collection/config/asset_config_file.txt

Figure 5.4 Asset details form

**Form 2: Asset Administration**

As data asset details are collected, this form aims to identify each asset administration details such as management IP address, port, and the required access conditions as illustrated in Figure 5.5.

**Asset Administration**

Please select Data Asset..

-----

**Data Assets Administration Details**

Mgmt. IP address and Port

Mgmt IP: \_\_\_\_\_ Mgmt Port: \_\_\_\_\_

Mgmt. Agent

Direct

Mgmt. Access Location

Direct Access

Mgmt. Privilage Required

USER

Save and Add New

Back Next

**Data Assets mgmt Table**

Below find assets mgmt list

#	Asset. ID(#)	Mgmt. IP	Mgmt. Port	Mgmt. Agent	Mgmt. Access	Mgmt. Privilage
1	Core-Switch-01 (Net1)	192.168.10.1	22	SSH-Terminal-01 (Ag2)	Local_Network	Admin

Figure 5.5 Asset administration form

**Form 3: Asset Dependency**

This is the last form in the infrastructure collection details; the form relies on complementing previously collected data with dependency details of the asset. The next stage uses the dependency information to build data flow diagrams. The form collects information about assets such as dependent assets, dependency type, and access requirements as illustrated in Figure 5.6.

Welcome Mohammed Alwaheidi (DA)  
This wizard's purpose is to understand infrastructure Dependency that supporting business services that are crucial to business services.

Please fill below inputs:

**Asset Dependency**  
Please select Data Asset..

**Data Asset Dependency Details**

Dependant Asset ID

Asset Dependency Type  
Exchange

Asset Dependency Access Requirements  
Direct Access

Save and Add New

Back Finish

**Data Assets Dependency Table**  
Below find assets Dependency list

#	Asset. ID(#)	Dep. Asset	Dep. Type	Dep. Access
1	Core-Switch-01 (Net1)	Edge-Router-01 (Net3)	Exchange	Local_Network

Figure 5.6 Asset dependency form

Overall, the three forms are designed to accommodate multiple iterations of information gathering pertaining to assets and dependent assets. Furthermore, the platform provides an important feature by providing an overview of collected data at each form in a table at the bottom of each page that can be revised once incorrect data is observed.

### 5.4.3 Threat Analysis Application

The Analysis process of the d-TM platform is a core component, where the analysis of collected data, weaknesses, and threats is conducted. Also, it includes the identification of suitable controls. This application consists of three sequential functions (aka modules): data analysis, threat analysis, and lastly threat mitigation. However, each of these modules is served by multiple supporting web pages.

On the other hand, the threat analysis application is accessed based on Role-Based Access Control (RBAC). This means that only users who have been assigned the "Analyst" role can access the threat analysis functions of the platform. This helps to ensure that only authorized users have

access to sensitive information. The threat analysis modules as presented in the platform are illustrated in Figure 5.7.

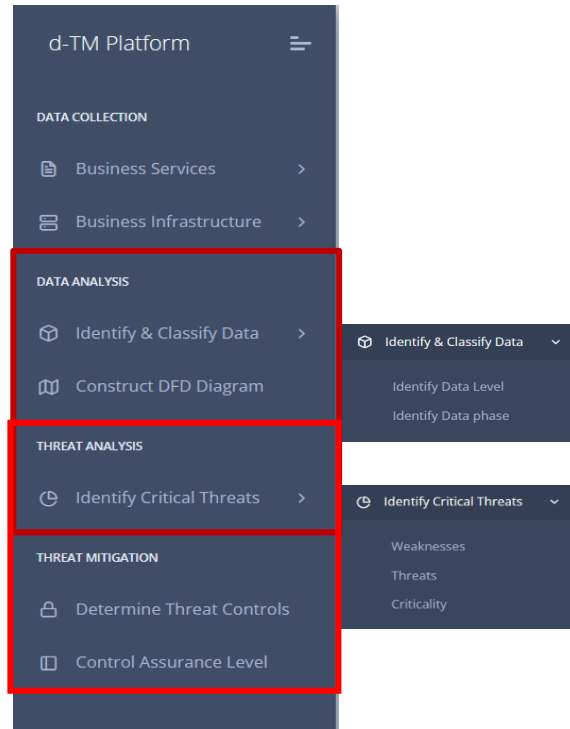


Figure 5.7 Threat Analysis Functions

## Module 1 Data Analysis

The Data Analysis module is a crucial tool for security analysts who want to understand the flow of data within an organization. By understanding the data flow, analysts can identify potential security risks and weaknesses. The d-TM platform analyses collected information about each asset, including identifying data levels and phases. This module also provides an important feature for constructing data flow diagrams (DFD) based on collected information about data assets.

To begin the data analysis process, the analyst needs to review each asset configuration to identify security-related information to data, and then assign appropriate data-related levels (management, control, or business). Once the data levels are identified, analysts need to assign the related data phase of each data level obtained from the data asset. The data analysis module concludes the outcome of the process in a table format that shows asset identification, data level, data phase, and related data asset configuration, as shown in Figure 5.8.

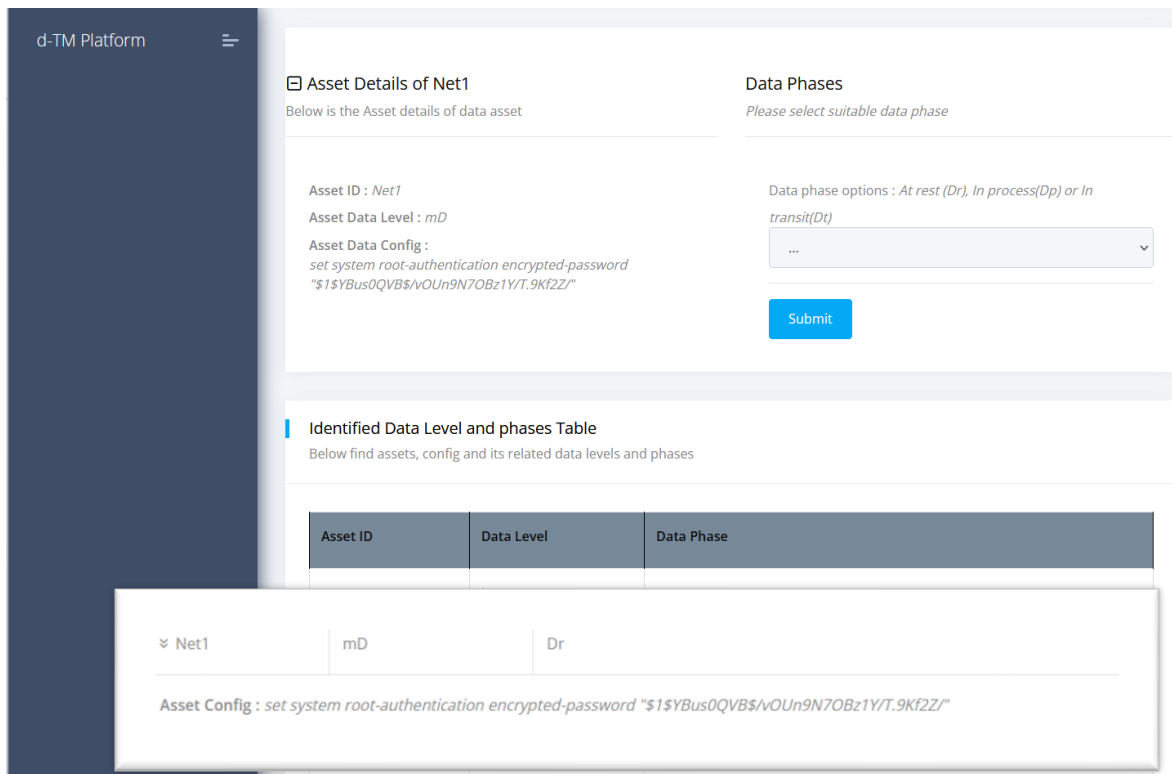


Figure 5.8 Data analysis interface

## Module 2 Threat Analysis

The threat analysis module is a critical component of the d-TM platform. This module conducts data threat analysis, which is a multi-step process that involves the following linked webpages:

- Identifying weaknesses: This webpage allows users to identify potential weaknesses in their data assets. These weaknesses can be technical, organizational, or operational in nature.
- Identifying threats: This webpage allows users to identify potential threats to their data assets. These threats can be natural, human-caused, or malicious in nature.
- Determining the threat's criticality: This webpage allows users to assess the criticality of each threat. This is done by considering the likelihood and impact of each threat.

- ***Weakness Identification Webpage***

The data asset weaknesses webpage is designed to support security analysts in identifying security weaknesses in data assets. The webpage lists all assets and their related information collected in previous stages and presents the information in a card-style format. The left side of each card contains the collected details, while the right side is designed for the security analyst to act.

Security weaknesses are imported to the platform from the Common Weakness Enumeration (CWE) catalogue. The security analyst must review the details of each asset and select the associated weaknesses for each asset card. Once the weaknesses have been selected from the dropdown menu, the Submit button is used to submit the data and associate the weaknesses with the asset in the database, as shown in Figure 5.9.

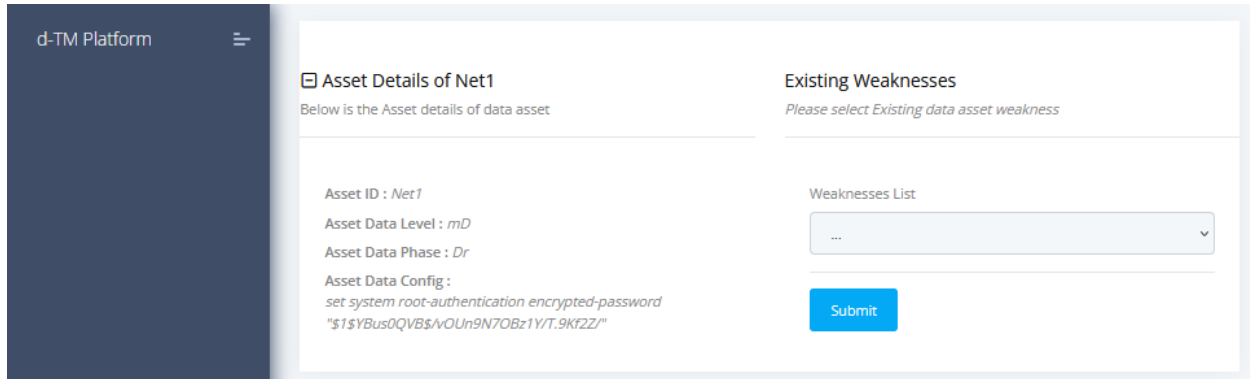


Figure 5.9 Weakness Identification Interface

The output of the process is presented at the bottom of the webpage in a table-style format. This table shows the assets associated with weaknesses, as well as other related details, such as the data type and the severity of the weakness. The table is shown in Figure 5.10.

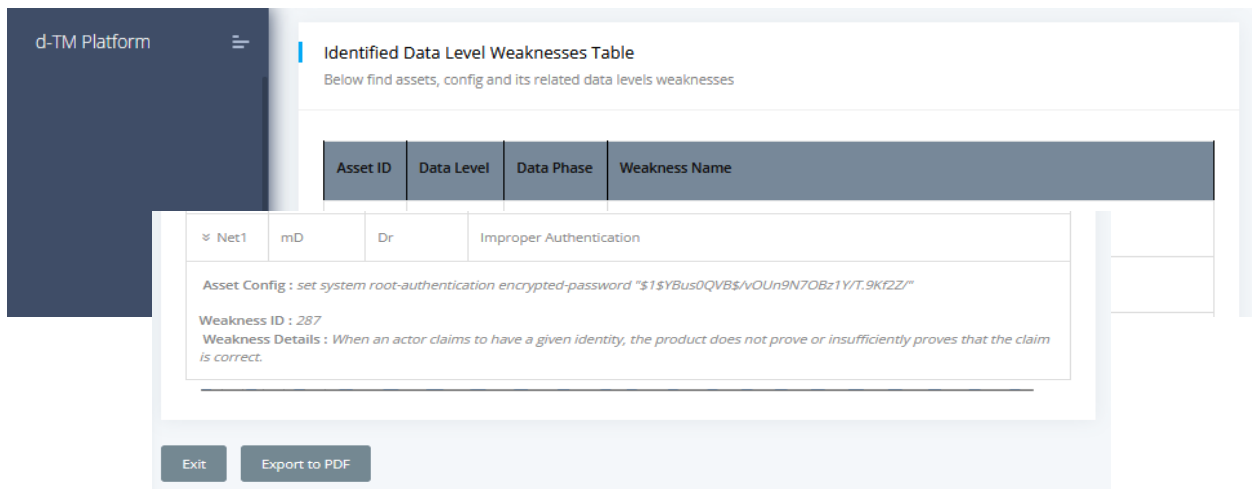


Figure 5.10 weakness identification table

- **Threat Identification Webpage**

The Threat Identification webpage is the second webpage in the process of threat analysis. This webpage retrieves the weaknesses, data, and asset details collected in the previous webpage to use in the process of threat identification. The webpage presents the information in a card-style format, with the collected details located on the left side of each asset card. The security analyst must



review the information on each asset card and select a suitable threat associated with each weakness from the dropdown menu on the right side of the card.

Security threats are imported to the platform from the Common Attack Pattern Enumerations and Classifications (CAPEC) catalogue. Once the threats have been determined and submitted to the database, the output of this process is presented at the bottom of the webpage in a table-style format, as shown in Figure 5.11.

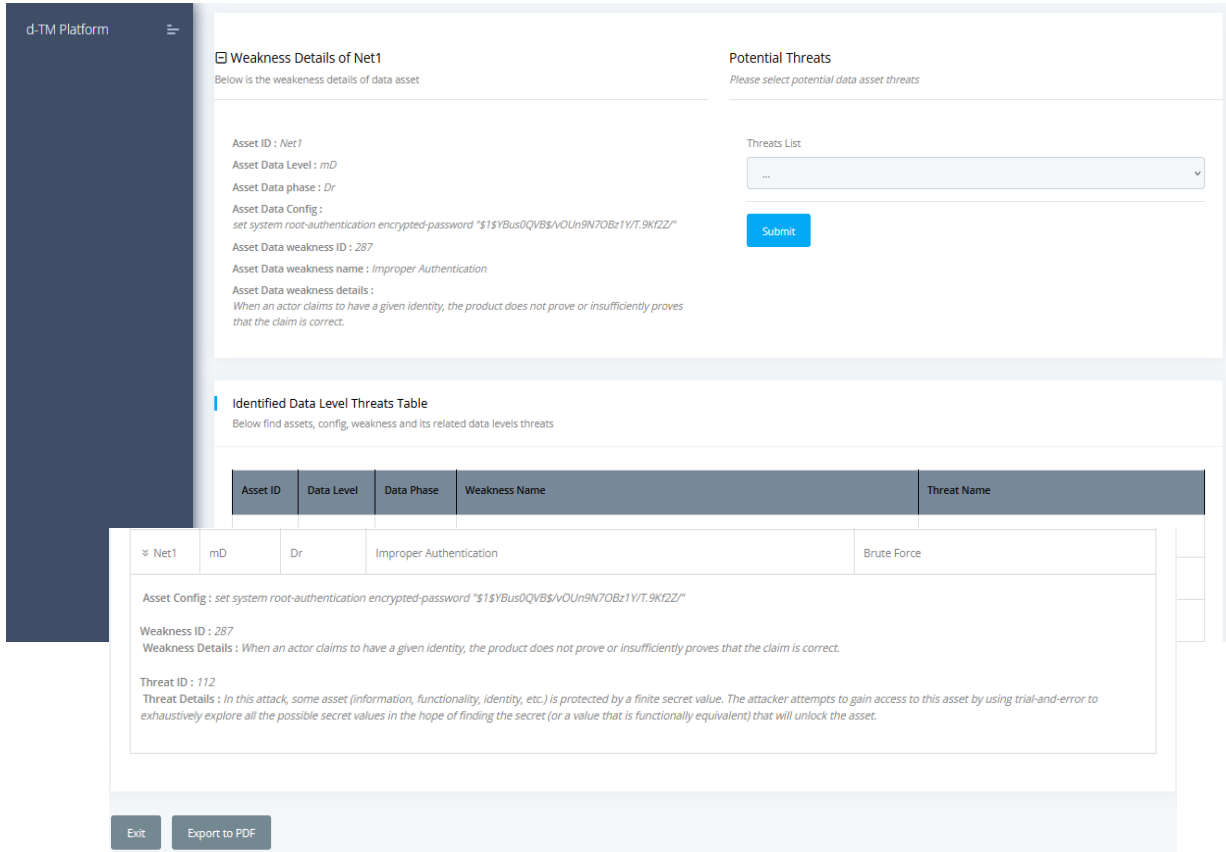


Figure 5.11 Threat identification interface.

- **Threat Criticality Identification Webpage**

The threat criticality webpage is the last webpage in the process of threat analysis. This webpage allows security analysts to evaluate the criticality of each threat to data assets. The webpage presents a list of d-TM threat criticality factors and metrics, and the security analyst must select the values that are most appropriate for each threat, as illustrated in Figure 5.12. The selected values are then used by the d-TM platform to calculate the overall criticality that can be used to prioritize the remediation of threats.

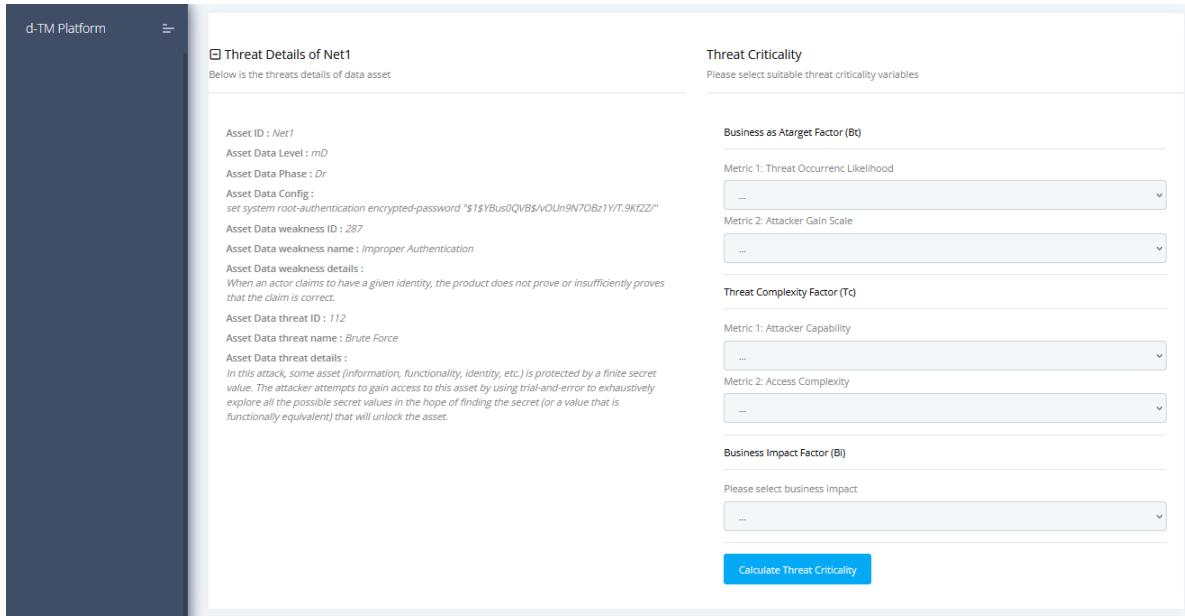


Figure 5.12 Threat criticality interface.

The ‘Threat Criticality Webpage’ is a crucial tool for security analysts who want to assess the risk of each threat to data. By understanding the criticality of each threat, analysts can make informed decisions about how to mitigate the risk. The criticality of each threat is presented in a colour-coded value based on threat criticality in the table produced at the bottom of the webpage, as shown in Figure 5.13.

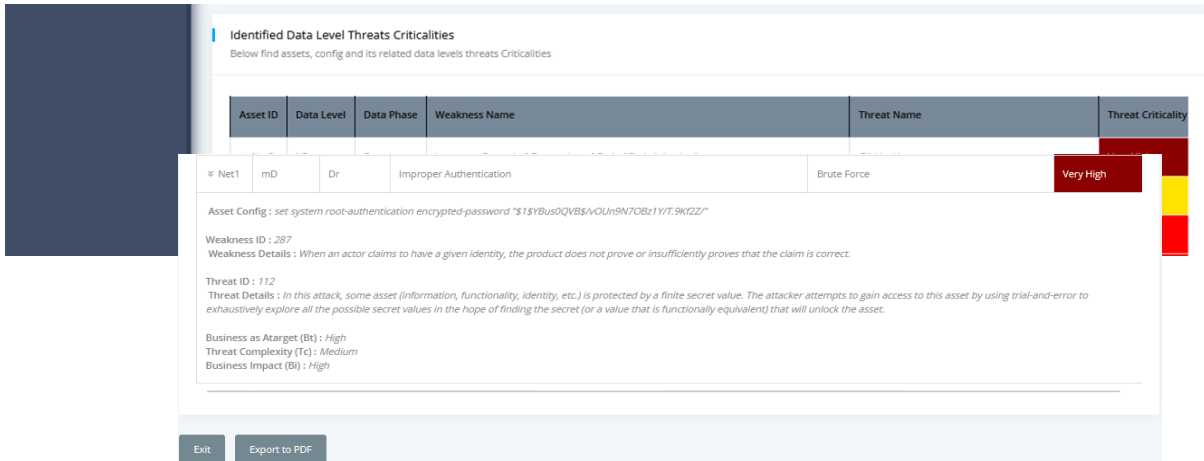


Figure 5.13 Threat criticality outcome table.

### Module 3 Threat Mitigation

The Threat Mitigation module is the third module in the threat analysis application of the d-TM platform. This module aims to extend the threat analysis capabilities of the platform by providing suitable controls and assurance levels. The threat mitigation module builds on the outcome of the

threat analysis module. The security threats and their criticality are imported from the threat analysis module. The threat mitigation module provides two webpages:

- **Determining Security Controls:** This webpage allows security analysts to select suitable security controls to mitigate the threats. The webpage presents a list of security controls, and the security analyst must select the controls that are most appropriate for each threat.
- **Calculating Assurance Level:** This webpage allows security analysts to calculate the assurance level of each control to the threat. The assurance level is a measure of the confidence that the control will be effective in mitigating the threat.

The threat mitigation module is a valuable tool for security analysts who want to mitigate the risk of threats to data. By selecting suitable security controls and calculating the assurance level of each control, analysts can reduce the risk of data breaches and other security incidents.

- **Security Controls Webpage**

This webpage retrieves all previous assessment details regarding data, weaknesses, threats, and criticality. The analyst will go over each threat and select a control that is appropriate for it. The National Institute of Standards and Technology (NIST) SP 800-53 Rev. 5 control list is used as a source for importing security controls. Finally, as illustrated in Figure 5.14, threats and controls are presented in the form of a table.

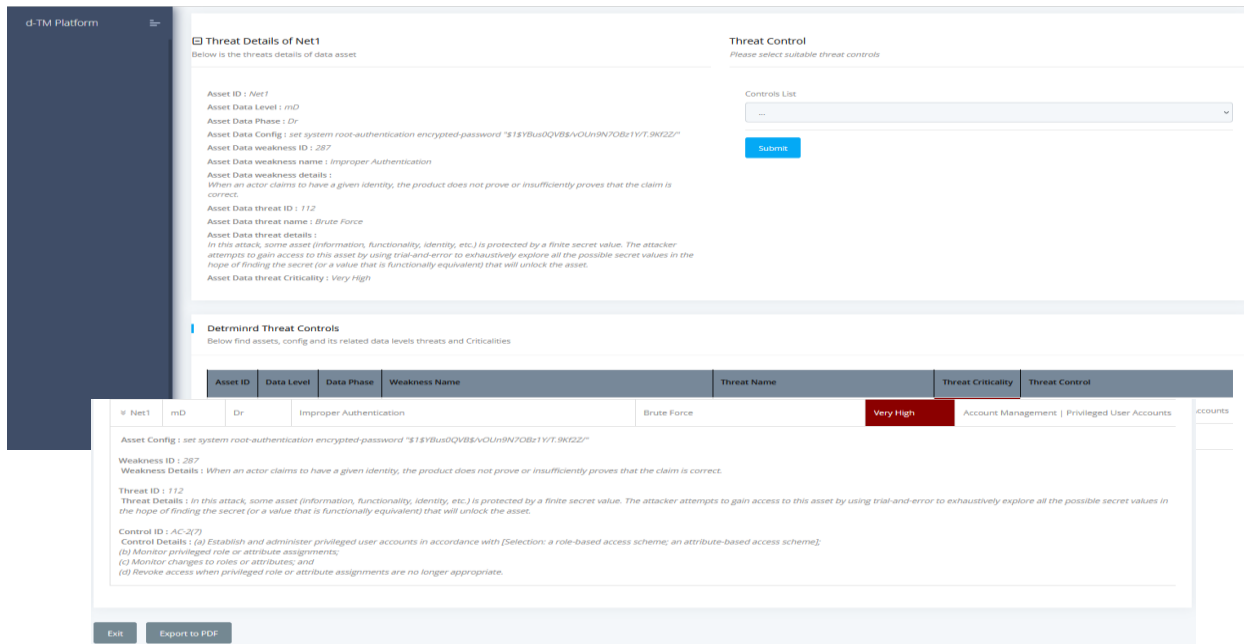


Figure 5.14 Threat mitigation interface.

- **Control Assurance Webpage**

At this point, the level of control assurance will be analysed and evaluated. While the controls are being determined, the webpage will assist security analysts in evaluating each control determined

based on the d-TM evaluation factors. The platform will calculate the level of assurance after the appropriate control assurance factors have been chosen, and then it will save this information in the database. As can be seen in Figure 5.15, the level of assurance that is associated with each control is detailed in the figure and in the table that is color-coded based on the assurance level.

The screenshot displays the 'd-TM Platform' interface. On the left, a sidebar shows 'Threat Details of Net1'. The main area is divided into two sections: 'Threat Details of data asset' and 'Threat Control Assurance'.

**Threat Details of data asset:**

- Asset ID : Net1
- Asset Data Level : mD
- Asset Data Phase : Dr
- Asset Data Config : set system root-authentication encrypted-password "\$1\$YBusDQVBS4OLh9N7Qbz1Y7t.9K2Z"
- Asset Data weakness ID : 287
- Asset Data weakness name : Improper Authentication
- Asset Data weakness details : When an actor claims to have a given identity, the product does not prove or insufficiently proves that the claim is correct.
- Asset Data threat ID : 112
- Asset Data threat name : Brute Force
- Asset Data threat details : In this attack, some asset (information, functionality, identity, etc.) is protected by a finite secret value. The attacker attempts to gain access to this asset by using trial-and-error to exhaustively explore all the possible secret values in the hope of finding the secret (or a value that is functionally equivalent) that will unlock the asset.
- Asset Data threat Criticality : Very High
- Data threat control ID : AC-2(7)
- Data threat control name : Account Management | Privileged User Accounts
- Data threat control details : (a) Establish and administer privileged user accounts in accordance with [Selection: a role-based access scheme, an attribute-based access scheme]; (b) Monitor privileged role or attribute assignments; (c) Monitor changes to roles or attributes; and (d) Revoke access when privileged role or attribute assignments are no longer appropriate.

**Threat Control Assurance:**

Please select suitable threat controls assurance factors

- Controls Completeness: [Dropdown menu]
- Controls Complexity: [Dropdown menu]
- Controls Effectiveness: [Dropdown menu]
- [Submit button]

**Identified Threats Control Assurance Level:**

Below find assets, config and its related data level threats controls assurance

Asset ID	Data Level	Data Phase	Weakness Name	Threat Name	Threat Criticality	Threat Control	Control Assurance
Net1	mD	Dr	Improper Authentication	Brute Force	Very High	Account Management   Privileged User Accounts	High

A detailed view of the selected threat and control is shown below the table, including details for the weakness, threat, and control, along with their respective assurance levels (Control Completeness: High, Control Complexity: Medium, Control Effectiveness: High).

Buttons for 'Exit' and 'Export to PDF' are located at the bottom left of the interface.

Figure 5.15 Control assurance interface.

The d-TM platform is also being developed to provide additional supporting features, such as a Dashboard and the ability to extract the outcomes of each process in a printable version called "PDF."

## 5.5 Conclusion

Cyber threats are increasing these days, attack surfaces are changing due to the continuous innovation in technology. Eventually, cyber technologies are running data, where data could be operational data or business data, also it can be found in three forms such as under processing, traversing over the network or steady in storage. Hence, these data levels and forms are subject to cyberthreats, so it needs to be secured.

d-TM focuses on data, where data levels and their status are the core principles of threat analysis. The model considers data as a source of analysis that bridges the research gap in the data-driven threat modelling knowledge domain, where some research is either old and lacking today's attack surfaces or focusing on business data only.

The research model focuses on treats targeting data directly or indirectly, where threats need to be identified and prioritized. Further, these threats are determined early by the model due to the concept of looking after the weakness that materializes the threat. Eventually, identified threats are addressed by suitable controls.

Lastly, the threats modelling process is empowered by a tool to automate the analysis process, which aims to enhance the time and effort to analyse threats targeting data.

# CHAPTER *SIX*: THE d-TM EVALUATION

## 6.1 Introduction

This chapter demonstrates the practicality and effectiveness of the proposed data-driven threat modelling (d-TM) approach, *three real-world case studies were conducted*. These case studies involved a *Solutions provider*, a *Cloud-enabled*, and a *Healthcare organization*. These organizations are looking to ensure their business revenue and operational experience through the adoption of robust security assurance measures. This chapter also provides an overview of the d-TM implementation process on real industrial case studies and presents observations derived from utilizing the d-TM approach to analyse threats in these scenarios. The main objectives of the evaluation are twofold: firstly, to assess the applicability and relevance of employing the d-TM approach, and secondly, to identify any potential challenges or issues that may arise during the implementation of the d-TM approach. Figure 6.1 illustrates an overview of the case scenarios and d-TM application used in this chapter.

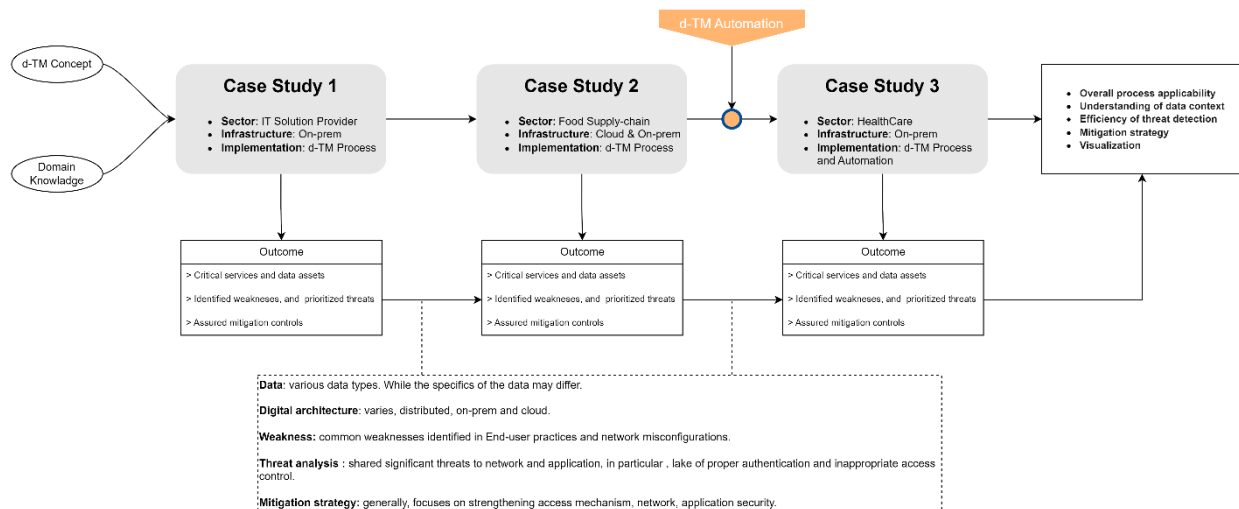


Figure 6.1. d-TM deployed case scenarios.

The case scenarios highlighted in the Figure represent different contexts from the industry, supply chain, service provider, and healthcare. The d-TM process is deployed manually (human efforts) in the first two cases, while the third case study outcome is produced using the d-TM automation platform. Each context represents different infrastructure deployments, such as remote, cloud, and on-prem deployments. The output of each case study concludes the threat analysis process of the d-TM model, including a detailed assessment of the data assets, weaknesses, potential threats, and suggested security controls to mitigate threats effectively. Furthermore, despite each organization's unique characteristics, the results across these case studies are compared to confirm the thesis objectives and identify similarities to threats and weaknesses faced by these organizations. This comparison helps to validate the effectiveness and practicality of the model using different real-case scenarios. However, demonstrating consistent results over these case studies provides evidence of validation that supports the thesis objective, which is around developing an effective data-driven threat analysis model for enhancing cybersecurity assurance.

The deployment of the d-TM methodology and tool in the case studies was characterized by a collaborative approach that began with an informative workshop for stakeholders, ensuring their

effective contribution to the process. Stakeholders were primarily involved in the initial phase, providing crucial business services and asset data. This phase set a solid foundation for the d-TM process. The deployment continued with a structured data collection process, involving extensive online meetings and secure communications to gather essential business and infrastructure information. The subsequent analysis, threat identification and mitigation phases were automatically conducted with the aid of the author - assuming a security expert. The automated d-TM was applied for the third case study; while the initial two cases were conducted manually. However, this evaluation demonstrates a successful integration of stakeholder engagement and Author in the application of the d-TM methodology in real-world case scenarios.

## **6.2 d-TM Application: A Real Industrial Case Study (Food Supply Chain)**

### **6.2.1. Case Study Scenario**

The case scenario revolves around a Middle Eastern fast-food chain restaurant, which was established in the early 1980s and has since grown to encompass over 300 locations with a workforce of 5,000 employees. The organization utilizes the SAP S/4HANA (on-premises) platform to manage various aspects of its operations, including restaurants, trucks, material orders, and supply chain processes. With the aim of enhancing infrastructure security and performance, minimizing downtime, and establishing a robust disaster recovery strategy, the company has decided to migrate its sales operations, hosted on the SAP platform, to the cloud. Specifically, the organization opts for the infrastructure-as-a-Service (IaaS) cloud model to gain more control and flexibility over its assets. By adopting S/4HANA on Google Cloud, the company experiences improved system stability, faster extraction of monthly financial reports, and a significant reduction in IT helpdesk inquiries. As the organization moves its critical infrastructure and data to the cloud, it becomes imperative to identify potential threats associated with this cloud migration and determine appropriate security solutions. Figure 6.2 provides an overview of the organization's overall architecture, which encompasses both internal and outsourced infrastructure. The sales operations are hosted on the Google Cloud platform and accessed by the sales team from local networks and restaurants. Each restaurant connects to a wireless network, which, in turn, connects to an internet modem. On the other hand, finance and management personnel access the cloud using the corporate internet, as do the corporate cloud administrators. Within the cloud platform, a virtual firewall instance serves as a cloud gateway for sales operation applications. The cloud application utilizes three cloud computing instances: SAP application NetWeaver, a database (DB), and storage. However, the remaining services are still housed in the local data center. The focus of this evaluation lies specifically on the sales operation that has been outsourced to the cloud, as it represents the organization's most critical service.





Process ID	Process Name and Description	Relevant Service ID	Service Name and Description	Service Criticality
<b>Bp0</b>	<b>Sales Operation</b> This process is critical to the company's business continuity, where sales operations are the main source of revenue for the company.	Bs0	<b>SAP Platform</b> This system provides a platform for sales representatives to do day-to-day sales tasks such as sales, checkout, balance, and purchasing. On the other hand, it provides management with information about all sales data and needs for planning and supporting the sales process.	H

<b>Bs0</b>		
<b>Asset details</b> <i>Asset<sub>id</sub> (Name/Role, Brand name, SW version)</i>	<b>Asset Administration</b> <i>Mgmt. (port, agent, access, privilege)</i>	<b>Asset Dependency</b> <i>Dep. (asset, type, access)</i>
Agt0(Web Browser, Google Chrome, 101.0.4951.54)	-	-
Agt1(SSH Terminal, Putty, 0.74)	-	-
Net0(PoS Internet Modem, ZNID, S3.1.135)	Mgmt. (443, Agt0, W/LAN, Local Admin)	Dep. (Internet link, Exchange, Direct)
Net3(Cloud vFW, Fortigate-VM, 5.4)	Mgmt. ((443, GCP-Console, GCP-Shell), Agt0, LAN, Security Admin group)	Dep. (GCP GW, Exchange, Local Network)
		Dep. (Jump-Server, Exchange, Local Network)
		Dep. (SAP Application, Exchange, Local Network)
		Dep. (GCP Cloud Portal, Control, Local Network)
Cmp0(Cloud Compute-1, SUSE Linux, 15)	Mgmt. ((22, GCP-Console, GCP-Shell), (Agt1, Agt0), Remote (Internet), GCP Admin group)	Dep. (SAP Application, Process, Direct)
		Dep. (Cloud vFW, Exchange, Local Network)
		Dep. (Cloud DNS, Process, Remote Network)
		Dep. (GCP Cloud Portal, Control, Local Network)
App0(SAP Application-1, NetWeaver, 7.5)	Mgmt. ((443, GCP-Console), Agt0, Remote (Internet), GCP Admin group)	Dep. (SAP DB, (Process, Store), Local Network)
		Dep. (Cloud Compute-1, Host, Direct)
		Dep. (GCP Cloud Portal, Control, Local Network)

Table 6.1. Table of critical business processes, services and assets.

### Activity 2 Data Analysis

This activity is the data analysis process of the d-TM; it begins once the information relating to the business process is gathered. It aims to identify the related data that supports the identified business process. The activity aims to analyse collected information that indicates how an asset

handles data at every level and phase. In addition to data presentation while moving across platforms. The identified actors are presented below:

- Business User (aka USR), which refers to any staff of sales, finance, or restaurant front-end representatives that are using Point-of-Sales machines. This actor has limited privileges and is enabled to access the SAP application for business-related activities such as registering or monitoring sales orders.
- Business Operator (aka OPR), which refers to any IT Staff working for the organization that access infrastructure or services for any administration activities such as troubleshooting. This actor has a high privilege that enables full control of the asset.
- Business System (aka SYS), which refers to any system-to-system relationships, such as SAP application and SAP DB. Also, CSP automation controllers and organization compute. It is a system or process exchanging data using API calls, for instance. The system could have full or restricted privileges, which depends on its role. For example, the CSP console has full privilege over organization compute hosted in the cloud.

Collected data at the previous stage is analysed by the team; the outcome is driven by the understanding of each asset, actor type, data level and phase. Table 6.2 depicts the Bs0. The first row represents the analysis outcome, while business users (USR) generate business data (bD) by accessing the service that is hosted on the cloud using the Chrome browser (Agt0) installed on the PoS machine; show data at each phase and how it's stored (Dr), processed (Dp) and transmitted (Dt). As a result, the browser is running default configuration and open-source plugins that are not used for business purposes, which shows no governance on the existing configuration to control software features such as storing sensitive data about sessions information locally with no protection or the presence of malicious code that interferes with the browser function such as intercept business data or escalate the attack to gain access to the local system. Eventually, this could lead to a threat to data while at rest (Dr) or processing (Dp). On the other hand, business data is sent over a wireless network that is also shared by the branch staff, who use their personal devices to access the Internet. The data, while transmitted (Dt) is at risk due to any compromised devices connected to the same network.

Service (Actor, Asset, data-level)	Data phases information
Bs0(USR,Agt0,bD)	(Dr) Browser saves session information locally, such as certificates, cookies, and history. Also, Data is stored in a single user profile for all employees. (Dp)Non-business Open-source plugins are installed. (Dt) Data is sent over wireless that is used for business and non-business purposes.
Bs0(USR,Net0,bD)	(Dt) PoS Modem is loaded with the default setting except for the wireless setting (WPA2 encryption). The admin page is open to access from any wireless SSID using HTTP. As well as locally stored credentials.
Bs0(USR,Net3,bD)	(Dt) Cloud vFW external interface eth0 is configured with a public IP address using a basic setting. Data volume is not restricted. Traffic is allowed based on any source to the SAP app using the IP address and port (443). However, the admin console is not accessed from the internet.

Bs <sub>0</sub> (USR, Ap <sub>0</sub> ,bD)	(Dp) SAP application is running as a system privilege. SAP application uses basic usernames and passwords for user authentication. SAP applications exchange data while processing over multiple ports with SAP DB. (Dt) SAP application data is sent over HTTPS to business users. SAP application data is forwarded to Net3 using private VN. SAP application authentication data is sent over LDAP protocol.
Bs <sub>0</sub> (OPR, Ag <sub>1</sub> ,mD)	(Dr) GCP computes, and cloud vFW identification information such as IP, port, and username is saved in the putty software for easy access. The private key file for cloud computing is stored locally on the system admin machine. (Dp) The putty software is installed as a trusted system-level process. The putty software does not require any authentication to run. (Dt) The putty software sends data over SSH or SFTP based on an IP address as an identifier.

Table 6.2. data level and phase analysis.

Figure 6.3 represents the business service (Bs<sub>0</sub>) data flow diagram that shows a business user using a PoS machine to access the sales operation application that is hosted in the cloud. The figure shows a salesperson using a web browser at a PoS machine to access cloud-hosted services, and the data is traversing over multiple assets till it reaches its destination. The figure represents the data levels and phases at each asset to facilitate sales operation application.

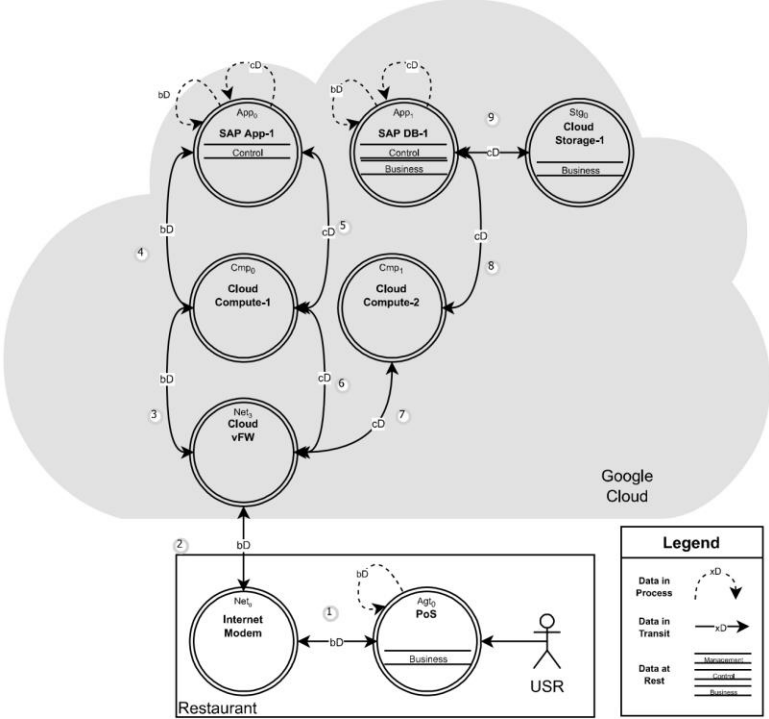


Figure 6.3. A data flow diagram of a business user accessing the SAP application.

**Activity 3 Threat Analysis**

This phase focuses on the threat analysis for the identified critical business service (Bs<sub>0</sub>). That includes identifying weaknesses, related data threats and criticality of identified threats to data.

The organization is turning to the cloud for agility and scale-on-demand as they modernize or convert their SAP systems (on-prem) to SAP S/4HANA(Cloud). Adding additional cloud services or managing hybrid environments expands an organization's threat surface. Security threats are aimed at the SAP Application, the online interface, and PoS devices that link to SAP. Cyber threats employ infrastructure as a point of entry to gain access to SAP's sensitive data. SAP does not typically give infrastructure security guidelines, and SAP's Security Baseline Template leaves these issues to the client to resolve(Fortinet, 2021). As a result, d-TM considers this gap and provides the necessary threat assessment to the SAP platform and its surrounding infrastructure.

The threat analysis process starts by identifying the weaknesses of the data by following the CWE reference model. A total of 9 weaknesses are identified, which are relevant to the scenario. The evaluation considers two infrastructure assets (Net0 and Net3), an SAP front-end application, and two different agent software (Agt0 and Agt1) that are used by organization users to access or operate cloud infrastructure. The d-TM approach starts assessing infrastructure from the user end, which is subject to many threats. For instance, the business user is using a web browser (Google Chrome) that has an asset ID (Agt0) to access the SAP application; Agt0 is assessed to identify weaknesses by evaluating the existing configuration, user practices and hosting operating system. As a result,

- The browser is running with the default configuration.
- Some commercial plugins are installed not related to business.
- A single profile is used by multiple users for accessing the windows operation system.
- The PoS machine is using an insecure wireless network; the network is accessed and shared with restaurant staff for personal use.
- The PoS machine data stored in the Windows operating system is accessible by restaurant staff.

Table 5.3, the first row, indicates the weaknesses that were discovered during the inspection of (Agt0). The weaknesses are primarily demonstrated as an uncontrolled asset, which provides a risk to all data phases (Dr,Dp,Dt) for business data level(bD), as well as management data(mD), if an administrator uses this system to access sensitive data or do maintenance. The first weakness demonstrates that the browser and the organization lack security control or policy to prevent or evaluate any program installed on the browser. Additionally, the information from browser sessions, history, and saved passwords are stored locally with no access restriction or encryption, thus putting stored data (Dr) at risk. Following the CAPEC reference model, the model determines corresponding threats to the data once the weaknesses have been detected. As shown in Table 8, there are a total of nine threats that are relevant to the scenario. In addition to threat identification, the process aims to determine the criticality of each identified threat. The criticality of each threat is determined by the d-TM process using the correlation of three factors (Bt, Tc, and Bi). For example, based on the organization's security engineer, an Adversary in the browser attack shown in Table 6.3 does not often occur, but it has been experienced by the organization previously at multiple restaurants. This attack could give an attacker a moderate gain by inspecting the traffic that is generated by the PoS machine about the sales activity of the branch. also, could lead to inspecting the salesman's credentials, which is considered to have a minimal impact on business

due to the restricted privilege assigned to PoS sales representative's accounts. However, this weakness could provide high gains to attackers and a high impact on the business if the browser is used by organization system admins. As a result of this information, considering the browser is used by the salesperson, not the admin, the overall criticality of this attack is Low(L), where Bt (M), Tc(M), and Bi(L). Lastly, the threat analysis of (Bs0) data is concluded as a threat profile table, which contains information on weaknesses, threats and their overall priority as determined by the d-TM threat analysis technique. Table 6.3 depicts the format of the threats profile table; it is formed of various columns. The columns aim to represent a list of assets evaluated and mapped to their weaknesses and introduced threats. Threats and weaknesses are listed with descriptions and identifiers in CWE and CAPEC catalogues. The reason for using catalogue identifiers rather than custom-generated ones is to provide an industry-understandable language and reference. Additionally, the table provides the organization with the criticality of each threat so that a mitigation strategy can be devised.

<b>Asset<sub>id</sub> (Data-level, Data-phase)</b>	<b>Weaknesses</b>	<b>Threats</b>	<b>Criticality (Bt, Tc, Bi)</b>
<b>Agt0 ((bD,mD), (Dr,Dp,Dt))</b>	CWE-494: Download of Code Without Integrity Check.	CAPEC-662: Adversary in the Browser (AiTB).	(M, M, L)=L
Agt0((bD,mD), (Dr))	CWE-921: Storage of Sensitive Data in a Mechanism without Access Control.	CAPEC-196: Session Credential Falsification through Forging.	(L, L, M)=L
Agt1(mD, Dr)	CWE-922: Insecure Storage of Sensitive Information.	CAPEC-529: Malware-Directed Internal Reconnaissance.	(M, M, H)=H
Net0(md, Dt)	CWE-319: Cleartext Transmission of Sensitive Information.	CAPEC-102: Session Side jacking.	(M, L, L)=VL
Net0(mD, Dp)	CWE-284: Improper Access Control.	CAPEC-1: Accessing Functionality Not Properly Constrained by ACLs.	(M, M, L)=L
Net3(mD, Dp)	CWE-308: Use of Single-factor Authentication.	CAPEC-151: Identity Spoofing.	(H, M, H)=VH
Net3(bD, Dt)	CWE-770: Allocation of Resources Without Limits or Throttling.	CAPEC-125: Flooding.	(H, H, H)=VH
App0((bD,mD), Dp)	CWE-308: Use of Single-factor Authentication.	CAPEC-151: Identity Spoofing.	(H, M, H)=VH
App0(bD, Dp)	CWE-20: Improper Input Validation.	CAPEC-63: Cross-Site Scripting (XSS).	(H, H, H)=VH

Table 6.3. Threat profile for Bs0.

## Activity 4 Threat Mitigation

This activity provides a mitigation course for identified critical threats. The threat is determined by the system's weaknesses, d-TM simply suggests evaluating CWE's suggested mitigation procedures to establish controls. d-TM perceives detected weaknesses as a source of threats that must be addressed. The "Scope" and "Impact" features, which are provided under the "Common consequences" section of every weakness identifier, could give needed information about any weakness. Moreover, CWE "Potential Mitigation" aids in the comprehension of the required mitigation plan. These details provide d-TM a better idea of what the attackers stand to gain with this attack and what kind of control they would need. At this stage, however, specialized expertise is required to correlate these data to appropriate NIST controls. As a result, security analysts must enumerate each relevant NIST security control family using CWE standards' scope, impact, and mitigation techniques. Table 6.4 represents the required mitigation controls based on the Bs0 threats profile. The table includes threats that require immediate attention according to their severity. However, the threats assigned a Very-high (VH) score are addressed at this phase. The table is focused on four threats to mitigate, which target cloud infrastructure elements (Net3 and App0). The threats can be concluded to organization data by account takeover, D/DoS, and application input misuse. As a result of understanding implies weaknesses, the controls should be within the scope of identity management, resource management, and system data integrity. By mapping these scopes to the NIST controls family, the NIST IA (Identification and Authentication), SC (System and Communication Protection), and SI (System and Information Integrity) controls family.

Lastly, the table summarizes the assurance level analysis of each control in relation to risks for a specific type of data. For example, the control (IA-2) is designed to protect management data (admin credentials) from a Credential spoofing attack caused by a lack of strong authentication requirements. A multi-factor authentication (MFA) solution is recommended by the control. The examination of this control reveals that the solution provides a significant mitigation match to a specific threat, which MFA considers to be exceptionally successful. MFA, on the other hand, is sometimes difficult to implement due to the necessity for integration with vital systems. The outcome of evaluating this control is an eight rating, indicating that the control assurance to the firm is extremely high.

<b>Asset<sub>id</sub></b> (Data-level, Data-phase)	<b>Threats</b> <b>(Criticality)</b>	<b>Controls</b>	<b>Assurance level</b> (Ct, Ef, Cx)=OAL
<b>Net3(mD, Dp)</b>	CAPEC-151 (VH)	<u>IA-2(1)- Identification and Authentication.</u>	(H, H, M) = 8 (H)
<b>Net3(bD, Dt)</b>	CAPEC-125 (VH)	<u>SC-5(3)- Denial-Of-Service Protection.</u>	(H, M, L) = 8(H)
<b>App0((bD,mD), Dp)</b>	CAPEC-151 (VH)	<u>IA-2(1)- Identification and Authentication.</u>	(H, H, M) = 8 (H)
<b>App0(bD, Dp)</b>	CAPEC-63 (VH)	<u>SI-10(5)- Information Input Validation.</u>	(H, H, L) = 9 (H)

Table 6.4. A table representing threat controls to protect (Bs0) data.

### 6.2.3 Discussion

The application of d-TM into the studied context is promising. This section provides our observation after implementing d-TM into the studied context.

#### d-TM Process

The process begins with an understanding of the business logic, which highlights the criticalities of existing systems to business continuity and narrows the efforts to low-priority business services. The analysis reveals that the main process for the business revenue is the sales operation, which is implemented on a cloud-enabled service platform. This important service is an SAP solution. In terms of business logic comprehension, the study used DFD diagrams to demonstrate the relationship between corporate cloud-enabled systems and company endpoint assets; the flow diagram is shown to ease comprehension of how the underlying technology interacts with data directly or indirectly. Furthermore, the connection between cloud systems and research case study components such as cloud consoles and cloud computes. The actor scenario described in DFD indicates that business users utilize PoS computers running a web browser as an agent to access the cloud service application (SAP), where the request is initiated by the agent, the data traverses over numerous network elements, locally and remotely (cloud), before being handed over to the business application hosted on cloud computing and leveraging cloud storage. DFD components for various actors are evaluated to extract and analyse data based on the d-TM concept. The goal of this assessment is to evaluate acquired data about running infrastructure and identify relevant information about how each asset operates with the data. The process of data collection and analysis is based on a security-oriented architecture and code review. The case-study asset evaluation is carried out based on the created DFD, which presents the information that could lead to a weakness in the asset, resulting in a data threat. For example, Agt0 and Net0 are two critical components for running business applications in any branch; they run with basic configuration and are not governed by the company. Furthermore, App0's authentication methodology for business users relies on usernames and passwords. Similarly, Net3 for operator access.

#### Threats and control

As previously indicated, the company is moving to the cloud to upgrade its on-premises systems to the cloud. Cyber-attackers use infrastructure as an entry point to the target's sensitive data. In contrast, the SAP application does not provide infrastructure security recommendations or a security baseline template, leaving these concerns to the customer to resolve (Fortinet, 2021). The threat analysis approach is designed to examine information and DFD diagrams derived from the data collection phase. The study gave a sample conclusion of detected weaknesses as follows:

- Assets: A total of five assets are investigated and regarded as significant components in the cloud data path. Data is created either by the business user at the PoS machine or by the system administrator using the corporate laptop on the journey to the cloud. The assets are Ag0, Ag1, Net0, Net3 and App0.
- Weaknesses: A total of nine weaknesses are determined out of the five assets. The nine weaknesses are investigated to reveal the implied threats to data.



Overall, the weaknesses revealed that the organization does not take the restaurants' endpoints and network seriously when it comes to security. As a result, endpoints are running an unregulated Chrome browser and operating system, putting company data at risk whether it is created or saved locally or in the cloud by salespeople. System administrators, on the other hand, share the same weaknesses as users who use uncontrolled browsers and support IT tools like "Putty." This could be used to store sensitive information regarding an organization's operational data, such as the address of its servers. System administrators' laptops. These weaknesses could have an impact on company data when it is in transit, in process, or at rest. An attacker might use the flaws to intercept, manipulate, or exfiltrate data in the browser. Furthermore, an attacker may employ a malicious plugin to obtain access to system files or conduct a browser MITM attack.

Regarding data threats, the assessment discovered multiple weaknesses that could be utilized by an attacker to impose threats to data while in the cloud. As a result of the d-TM approach, there are multiple threats discovered in reference to identified weaknesses.

- Threats: A total of nine threats are identified out of five assets assessed. The threats are targeting data at infrastructure components in different layers such as agent, network, and application.
- Critical threats: A total of four threats are categorized as critical out of nine identified threats. The four threats need immediate attention in the case study. Denial of service, XSS, and social engineering are critical threats due to weaknesses in two assets (Net3, App0).
- Data: A total of two cloud assets out of five represent a high risk to data, where threats are targeting mainly two data levels, which are management and business data levels. Furthermore, the data is at risk in two phases, which are at transit and during the process.

The research has revealed that there are multiple possible threats within the cloud infrastructure, such as virtual cloud instances as application front-end interfaces, namely Net3. This device is configured to accept any request and has no means of controlling the volume of incoming traffic. This could be due to a constraint on the asset's capabilities or the assumption of the presence of a third-party security solution. In our scenario, no solution is available. As a result, D/DoS attacks are possible on these internet-facing assets. Cloud business services could be severely disrupted because of such an attack. This issue has been rated as an extremely high threat by the d-TM, and it must be addressed as quickly as possible. Furthermore, the front-end cloud application lacks a validation capability for incoming requests, making the application vulnerable to a Cross-Site Scripting (XSS) attack. This attack was also a high-ranking one. Finally, cloud apps and network assets rely on credentials that users have memorized; this weakness could lead to a variety of threats aimed at users, such as social engineering attacks. On the other hand, if a business user credential is taken, the impact is limited; however, this is not the case if the credential belongs to a system administrator. Nonetheless, when the system administrator is an actor, this threat is rated as extremely high to management data.

The d-TM model provides the organization with appropriate controls to help mitigate identified critical threats.

- Controls: Four controls are identified to mitigate the critical threats. The controls are mainly looking after authentication improvement, input validation and service assurance.

The organization is mostly under attack from threats that target system identification mechanisms, resource capacity management, and application input control, according to the high critical threats presented in the threat profile. As a result, the model suggested a set of measures to help organizations deal with the threats, such as implementing multi-factor authentication mechanisms, implementing a validation control on user inputs, and obtaining technology to monitor and respond to resource misuse.

### **6.3 d-TM Application: A Real Industrial Case Study (Service-provider)**

The proposed d-TM is implemented into a real industrial case study. The purpose of this evaluation is to determine the applicability of the d-TM in a real-world scenario. The d-TM process is systematically applied based on its activities and steps, and various data is analysed to determine its validity. A detailed description of the uses case scenario is presented and followed by the implementation.

#### **6.3.1 Case Study Scenario**

TSS is a leading cybersecurity services provider located in the Middle East, providing services to banks, government agencies, and commercial customers. TSS offers cybersecurity services such as comprehensive cybersecurity solutions, training, managed services, professional services, and cybersecurity consultation. The organization is operated by many functions, such as sales, marketing, human resources, information technology, professional services, consulting services, training, finance, and logistics. It has four offices in various locations, 90+ employees, and a revenue of 13 million dollars each year.

TSS is regarded as a managed security service provider (MSSP) for organizations seeking to outsource the monitoring and management of their systems. TSS's primary business is this service, which is built on a centralized Security Information and Event Management platform (SIEM) and other supporting technologies. This platform serves as a single cybersecurity monitoring solution, providing all cybersecurity intelligence for managed client devices. In addition, it gives a detection engine for any attack or breach of security policy. The TSS security analysts team handles the monitoring and detection platform, with management and incident response. Data is a vital asset for TSS, and as a cybersecurity solution provider, they are constantly searching for ways to analyse their data architecture against threats and provide superior protection for their and their customers' data assets. Figure 6.4 represents the TSS infrastructure topology that serves MSS service. The infrastructure is designed in a multilayer approach, where services are connected to network devices. The devices facilitate access to users, customer sites and the internet. Traffic in both directions to/from services, internet, wan, and customer sites is protected by firewalls. MSS service has two layers: front-end (customer-facing portal) and back-end. Customers are accessing MSS service using a customer-facing portal hosted in the DMZ zone and protected by the perimeter firewall; similarly, the DMZ zone is hosting data collection of customer devices. The TSS analyst

team is accessing back-end services using HTTPS and SSH for troubleshooting. However, the back-end services are protected by an internal firewall.

TSS continues to expand its business; as a result, it expands its infrastructure to accommodate client devices connected to the internet. Due to increased demand and the nature of the managed services TSS offers for its clients, TSS is seeing an increase in the number of cyberattacks targeting its services, particularly those connected to its MSS platform. TSS was compromised by multiple attacks targeting its infrastructure supporting the Managed Security Service (MSS) platform. DDOS attacks and Brute force attacks continuously target MSS services; MSS admins and analysts have also targeted victims for account takeover attacks. Nonetheless, TSS is an Anonymous name provided due to the confidentiality of the use case.

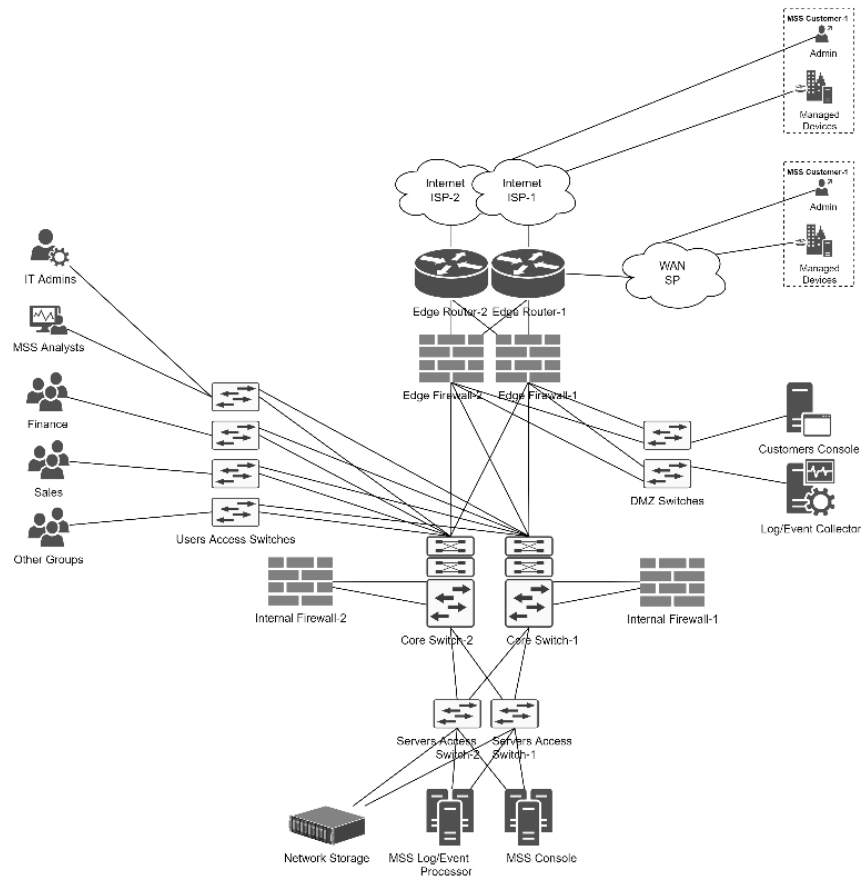


Figure 6.4. The TSS infrastructure topology.

### 6.3.2 d-TM Implementation

The following section outlines the process flow and outcomes derived from the implementation of the d-TM model in the case study. For this implementation, a dedicated team was constituted, comprising of three employees from the organization. The team tasked with executing this activity includes the sales director, MSS manager, and Senior SoC analyst, collectively forming the Threat Assessment and Security Strategy (TSS) team.

## Activity 1 Data Collection

The first component shows the important TSS business services and underlying infrastructure. The outcome of this component is summarized in the three tables below. Table 6.6 shows the results of understanding business context and priorities, and the focus for the next steps is on MSS business service. Where B stands for a business context, Bp stands for business process, and Bs is Business service.

Process ID	Process Name and Description	Relevant Service ID	Service Name and Description	Service Criticality
<b>Bp<sub>0</sub></b>	Sales and Marketing  <i>This process aims to provide TSS with all the requirements to support sales and marketing functions.</i>	Bs <sub>0</sub>	Sales and Deal Registration Service  This system provides a service to register sales opportunities determined by the sales team or business partners; the system provides information about running or forecasted sales deals such as sales stage, gain amount, close date, challenges, etc.	M
		Bs <sub>1</sub>	Marketing Service  This system provides the organization with a platform to demonstrate the organization's portfolio and announce new services, feeds, marketing campaigns, and contact information to customers visiting the system. Also, get customer inquiries, feedback, suggestions, etc.	NC
<b>Bp<sub>1</sub></b>	Human Resources  <i>This process aims to provide TSS with all the functions related to workforce operations such hire, retirement, compensations, etc.</i>	Bs <sub>2</sub>	HR Service  This system provides organizations with a platform to operate organization employee requirements, i.e., employee details, contract, time-off, payroll, hire, resignation, etc.	L
<b>Bp<sub>2</sub></b>	Finance  <i>This process aims to provide TSS with all the functions related to finance, including procurements, payments, logistics, salaries, ...etc.</i>	Bs <sub>3</sub>	Payroll Service  This system provides the organization with a platform to manage employee's payroll.	L
		Bs <sub>4</sub>	Order Management Service  This system provides the organization with a platform to manage and operate procurements, sales payment orders, purchase orders, and other business payments, i.e., sales commission.	M

<b>Bp3</b>	Core Business  <i>This process aims to provide TSS with all the Functions related to providing security services to customers, including training, managed services, solutions, and consultancy.</i>	Bs5	Managed Security Service (MSS) Service  This system provides the organization with a platform to monitor and operate customer security assets. This system is considered as a core business for the organization. It facilitates access to customer's devices and gets all required configuration, logs, and events that are used for security analysis.	H
		Bs6	Training and Education Service  This system provides the organization with a platform to serve customers who apply for available training and exams by the organization. Also, this system is used by internal employees for awareness sessions.	L
<b>Bp4</b>	IT Support Business  <i>This process aims to provide TSS with all the Functions related to IT infrastructure and services supporting day-to-day business, including Email, File Storage, Collaboration, etc.</i>	Bs7	Email Service  This system provides the organization with a platform to communicate internally and externally with customers by sending/receiving proposals, offers, feedback, etc.	H
		Bs8	Collaboration Service  This system provides the organization with a platform to communicate internally and externally by voice/video, i.e., conferences.	M
		Bs9	Shared Storage Service  This system provides the organization with network storage to store and exchange files, i.e., shared drives and folders.	L

Table 6.6. Business services and processes table

In reference to the above table, TSS is considering two services that are important to running the business. The scope of analysis below focuses on MSS service, which has Bs id ‘5’ and is ranked as a critical service to TSS. Table 6.7 shows the infrastructure components of the MSS service as well as technical specifications. The table provides a sample list of assets, as well as assets detected by d-TM. Later stages concentrate on the MSS Customer console application, which TSS customers use to analyse their security requirements.

Bs <sub>5</sub>		
<b>Asset details</b> <i>Asset<sub>i,d</sub> (Name/Role, Brand name, SW version)</i>	<b>Asset Administration</b> <i>Mgmt. (port, agent, access, privilege)</i>	<b>Asset Dependency</b> <i>Dep. (asset, type, access)</i>
Ag0(Web Browser, Mozilla Firefox, 84.0.2)	-	-
Ag1(Web Browser, Google Chrome, 94.0.4)	-	-
Ag2(SSH Terminal, Putty, 0.6)	-	-
Net0(Edge Router-1, CISCO ASR 1001, IOS 16.10)	Mgmt. (22, Ag <sub>1</sub> , LAN & Direct, Local Admin)	Dep. (Edge Router-2, Exchange & Control, Local Network)
		Dep. (Edge Firewall-1 & 2, Exchange, Local Network)
		Dep. (Internet ISP-1, Exchange, Remote Network)
		Dep. (MSS Event Processor Server -1,2&3, Storage, Local Network)
Comp0(MSS Customer Console Server-1, Lenovo, RHEL 7.5)	Mgmt. (22, Ag <sub>1</sub> , LAN and Direct, Local Admin)	Dep.(MSS Console Server-1, 2 & 3, Exchange, Local Network)
		Dep.(MSS Event Processor Server-1, 2 & 3, Exchange, Local Network)
		Dep.(DMZ Switch-1 &2, Exchange, Local Network)
		Dep.(Edge Firewall-1 & 2, Exchange and Process, Local Network)
		Dep.(MSS Customer Console App-1, Host, Direct)
		Dep.(DNS Server-1, Process, Local Network)
App0(MSS Customer Console App-1, IBM Qradar, 7.2)	Mgmt.(443, Ag <sub>0</sub> & Ag <sub>1</sub> , LAN (Ag <sub>0</sub> & Ag <sub>1</sub> ) and Direct (Ag <sub>1</sub> ), Local Admin)	Dep.(MSS Customer Console App-1 Process & Control, Local Network)
		Dep.(MSS Console App-2 &1, Process & Control, Local Network)
		Dep.(MSS Event Processor App-1, 2 &3, Process & Control, Local Network)
		Dep.(DNS Service App-1, Process, Local Network)
Strg0(MSS Network Storage-1, NetApp, 4.0)	Mgmt.(22 & 443, Ag <sub>0</sub> & Ag <sub>1</sub> , LAN (Ag <sub>0</sub> & Ag <sub>1</sub> ) and Direct (Ag <sub>1</sub> ), Local Admin)	Dep.(MSS Console App-1, 2 &3, Exchange, Local Network)
		Dep.(MSS Event Processor App-1, 2 &3, Exchange, Local Network)
		Dep.(Servers Access Switch-1 &2, Exchange, Local Network)

Table 6.7. Infrastructure details of MSS service.

## Activity 2 Data Analysis

Once the information relating to the business service is gathered, it is necessary to identify the related data that support the identified business process. This activity aims to discover useful information that indicates how an asset handles data at every level and phase. In addition to data presentation while moving across platforms. The identified actors are presented below:

- Business User (aka UR), which refers to any customer representatives that is using the MSS console. This actor has a limited privilege to view and monitor their managed devices' status and generate reports and support tickets.
- Business Operator (aka OPR), which refers to any TSS SoC analyst working for the organization to access infrastructure or services for any administration activities such as troubleshooting and incident response. This actor has a high privilege that enables full control of the asset.
- Business System (aka SYS), which refers to any system-to-system relationships, such as MSS applications and customer-managed devices. Also, systems intercommunications among MSS service components. However, the system could have full or restricted privileges, which depends on its role.

Collected data at the previous stage is analysed, and the outcome is driven by the understanding of each asset, actor type, data level and phase. Table 6.8 summarises the results of data analysis and observation of each identified TSS MSS service component.

<b>Service</b>  <i>(Actor, Asset, data-level)</i>	Data phases information
Bs <sub>5</sub> (USR,Agt0,bD)	(Dr) Browser using default defined configuration, saves session information locally such as certificates, cookies, and history.
Bs <sub>5</sub> (USR,Net0,bD)	(Dt) Router external interfaces are configured as basic as IP address, subnet and accept any traffic with full capacity.
Bs <sub>5</sub> (OPR,Agt1,mD)	(Dr) Browser using default defined configuration, saves session information locally such as certificates, cookies, and history. (Dp) Non-business Open-source plugins are installed.
Bs <sub>5</sub> (OPR,Agt2,mD)	(Dr) Agent has saved credentials for quick access. (Dr) Agent has no authentication mechanism. (Dp) Non-business Open-source plugins are installed.
Bs <sub>5</sub> (SYS,Net0,cD)	(Dt) Router uses BGP protocol with a basic IP address, subnet for authentication mechanism to peer router. (Dp) Server is using TSS DNS server for name resolution, and public NTP server IP for time sync.
Bs <sub>5</sub> (OPR,Net0,mD)	(Dp) Router is accessed using a basic authentication mechanism of username/password while DB is stored locally. (Dp) Router management interface is accessed from anywhere in the LAN using an In-band management interface. (Dp) Multiple credentials is created with full privilege.
Bs <sub>5</sub> (OPR,Cmp7,mD)	(Dp) Router is accessed using a basic authentication mechanism of username/password while DB is stored locally.
Bs <sub>5</sub> (SYS,Cmp7,cD)	(Dp) Server is using TSS DNS server for name resolution and a public NTP server IP for time sync.
Bs <sub>5</sub> (USR,APP0,bD)	(Dt) Application uses TLS 1.2 for encryption. (Dp) Application authenticates users using a basic authentication mechanism of username and password.
Bs <sub>5</sub> (OPR,APP0,mD)	(Dp) Application authenticates OPR using basic authentication mechanism of username and password. (Dr) Admin credentials are stored in the local DB for access authentication and authorisation

Table 6.8. Data-levels and phases analysis.

Following the conclusion of the analysis of data assets that support vital services, a data flow diagram is produced. Figure 6.5 depicts the business service (Bs5) data flow diagram, which demonstrates a business user scenario utilizing MSS service.



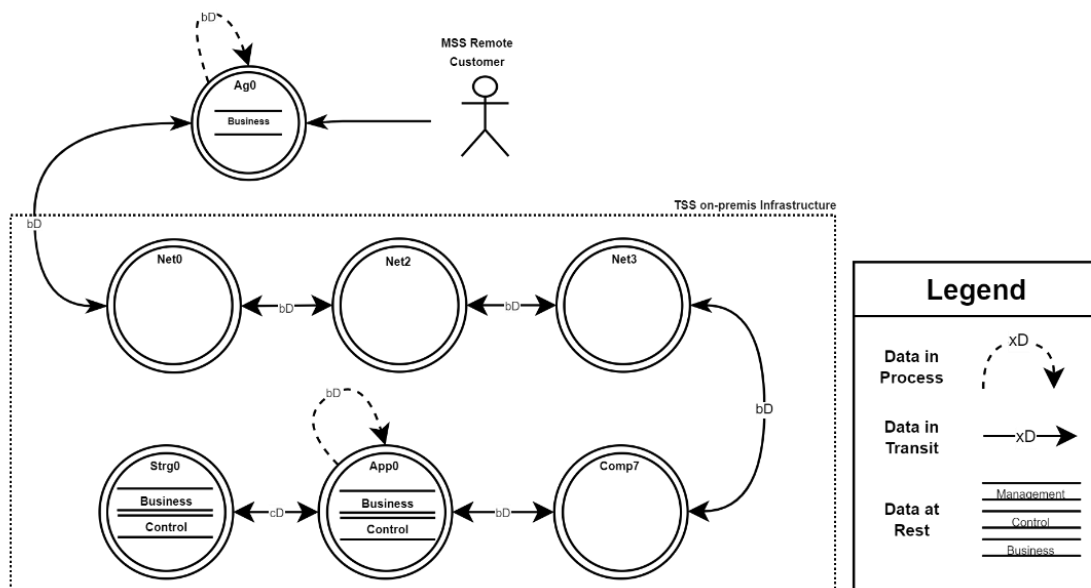


Figure 6.5. MSS service – Business Data-flow

### Activity 3 Threat Analysis

During this activity, the threat analysis for the identified key business service (Bs5) is the primary emphasis. This covers the identification of data assets' security flaws, associated threats, and the criticality of detected data security threats.

The CWE reference model is used as a starting point for the process of threat analysis, which begins with the identification of the data asset weaknesses. As can be seen in Table 6.9, a total of seven different weaknesses that are pertinent to the scenario have been discovered. The evaluation considers five infrastructure assets known as Ag0, Ag1, Ag2, Net0, and App0. Ag0, 1 and 2 are the components that allow customers and operators to access MSS services. In addition, Net0 is an essential component for gaining access to the hosted services provided by TSS through the internet. App0 is the MSS portal that is utilized to run customer-managed devices. Customers also use App0 for monitoring and reporting. The concerns detected during the examination of the assessed component are shown in the table below. The first weaknesses are mostly shown using unprotected software to access the service. It poses a threat to all data phases for business data level (bD) while using Ag0 and management data level (mD) for Ag1 and Ag2, where administrators utilize these tools to access TSS assets and extract sensitive data or perform maintenance.

Once the weaknesses have been identified, the CAPEC reference model is used to determine the possible threats to the data. Table 6.9 shows that there are a total of nine threats relevant to the scenario. The first rows of the table show the detected vulnerability, which is connected to an uncontrolled browser that stores sensitive data insecurely and enables any program (aka plugin) to be installed without first properly analysing the code's origin and integrity. This program may include malicious code that allows an attacker to tamper with browser functionality by exfiltrating,

manipulating, or disrupting processed/stored/transmitted data. These attacks are known as "Session Credential Falsification" and "Advisory in the Browser (AiTB)".

In addition to identifying threats, the procedure attempts to determine the criticality of each detected threat. The d-TM approach uses the correlation of three elements to determine the criticality of each threat (Bt, Tc, and Bi). The format of the threats profile table is given in Table 6.9; it is composed of several columns. The columns seek to reflect a list of assets appraised and mapped to their weaknesses and possible threats. Threats and weaknesses are catalogued with descriptions and IDs to CWE and CAPEC. The table advises the company about the criticality of each threat, allowing it to design a mitigation strategy.

<b>Asset<sub>id</sub> (Data-level, Data-phase)</b>	<b>Weaknesses</b>	<b>Threats</b>	<b>Criticality (Bt, Tc, Bi)</b>
<b>Agt0 ((bD), (Dr, Dp, Dt))</b>	CWE-921: Storage of Sensitive Data in a Mechanism without Access Control.	CAPEC-196: Session Credential Falsification through Forging.	(L, L, M)=L
<b>Agt1 ((mD), (Dr, Dp))</b>	CWE-494: Download of Code Without Integrity Check.	CAPEC-662: Adversary in the Browser (AiTB).	(M, M, L)=L
<b>Agt2(mD, Dr)</b>	CWE-922: Insecure Storage of Sensitive Information.	CAPEC-529: Malware-Directed Internal Reconnaissance.	(M, M, H)=H
<b>Net0(mD, Dp)</b>	CWE-284: Improper Access Control.	CAPEC-1: Accessing Functionality Not Properly Constrained by ACLs.	(H, M, M)=H
<b>Net0(mD, Dp)</b>	CWE-308: Use of Single-factor Authentication.	CAPEC-151: Identity Spoofing.	(H, M, H)=VH
<b>Net0(bD, Dt)</b>	CWE-770: Allocation of Resources Without Limits or Throttling.	CAPEC-125: Flooding.	(H, H, H)=VH
<b>App0((bD, mD), Dp)</b>	CWE-308: Use of Single-factor Authentication.	CAPEC-151: Identity Spoofing.	(H, M, H)=VH

Table 6.9. Threat analysis of MSS Bs5.

#### Activity 4: Threat Mitigation

This activity is the final activity in the assessment process. Hence, TSS MSS service cyber threats, weaknesses, and its priorities for business continuity are determined by previous activity. d-TM at this activity focuses on determining suitable controls that help the management team build a mitigation plan to address underlying threats. Also, the assurance level of each determined control to the organization security posture. Following the threats profile provided in past activity, security analysts must enumerate each applicable NIST security control family utilizing the scope, impact, and mitigation methodologies of CWE standards. Table 6.10 shows the necessary mitigation controls based on the Bs5 threat profile. According to their seriousness, the threats in the table

demanded quick intervention. Threats with a very high (VH) score, on the other hand, are handled at this phase. The table focuses on three mitigation threats that target MSS infrastructure components (Net0 and App0). Account takeover, denial of service, and asset manipulation are examples of risks to organizational data. Finally, the table highlights each control's assurance level analysis in connection to threats for a given type of data.

<b>Asset<sub>id</sub> (Data-level, Data-phase)</b>	<b>Threats (Criticality)</b>	<b>Controls</b>	<b>Assurance level (Ct, Ef, Cx)=OAL</b>
<b>Net0(mD, Dp)</b>	CAPEC-151 (VH)	<u>IA-2(1)- Identification and Authentication.</u>	(H, H, M) = 8 (H)
<b>Net0(bD, Dt)</b>	CAPEC-125 (VH)	<u>SC-5(3)- Denial-Of-Service Protection.</u>	(H, M, L) = 8(H)
<b>App0((bD,mD), Dp)</b>	CAPEC-151 (VH)	<u>IA-2(1)- Identification and Authentication.</u>	(H, H, M) = 8 (H)

Table 6.10. Threat mitigation and assurance of MSS Bs5.

## **6.4 Automation Evaluation of d-TM model: A Real Industrial Case Study (Healthcare)**

This case study is implemented using the d-TM automation platform, unlike the previous two cases where the d-TM process is manually implemented using human efforts. This case study aims to show the applicability of the d-TM automation tool and identify the added value to the d-TM application and challenges where applicable.

The aims of the evaluation:

- Feasibility of adopting the d-TM automated tool.
- Coverage of automated threat analysis and management.
- Identify weaknesses and associated critical threats.
- Determine assured threat mitigation strategy.
- Consolidation of expert opinion about d-TM and its tool

### **6.4.1. Case Study Scenario**

The case study is a healthcare provider(Hospital) located in Saudi Arabia. The hospital is ranked as one of the leading hospitals in the healthcare sector in 2021. Due to confidentiality reasons, we could not disclose the name of the hospital in this research. The hospital's main goal is to deliver patient care that ensures safety and maintains a level of quality. The hospital has become widely recognized as a leading institution, in Saudi Arabia and the wider Middle East region. The hospital is renowned for its healthcare services in fields such as cancer treatment organ transplants, heart diseases, neurological conditions, and genetic disorders. It plays a role as a centre for patients from both the country and the surrounding region.

#### **Healthcare Services**

This case study extensively explores the transformation journey of a hospital that offers both inpatient and outpatient medical care. As a leading institution in the field, the hospital has truly embraced electronic services to provide exceptional patient care and simplify administrative tasks. These services include the following:

- **General e-Services:** At the forefront of its digital initiatives, the hospital's general e-services span across four critical categories. From a certificate verification system, ensuring the authenticity of patients' medical credentials, to a dedicated platform for health education, the hospital endeavours to disseminate critical healthcare knowledge.
- **Patient e-Services:** Tailored to empower patients, the dedicated e-service portal serves as a digital gateway. Patients can effortlessly reschedule or cancel appointments, access comprehensive medical histories, retrieve previous laboratory test results, immunization records, medication prescriptions, and much more. This 24/7 digital assistance proves invaluable, especially during emergencies, rendering services at the patient's convenience.

- **Suppliers e-Services:** Fostering transparency and efficacy, the hospital's e-services for suppliers streamline the procurement process. With pivotal services like Bid Management and Supplier management, suppliers navigate a transparent bidding process, facilitated by the government's procurement initiative portal. This advanced-grade scheme adheres to stringent guidelines, assuring an equitable and meticulous bidding environment.
- **Telehealth-EMS Services:** Embracing the potential of virtual healthcare, the hospital's Tele-Emergency Medical Services (Tele-EMS) offers a specialized portal, enabling the Hospital's consultants to virtually collaborate with the EMS Department, addressing critical emergency cases with prompt expertise.
- **Employee e-Services:** Recognizing the pivotal role of its workforce, the hospital's e-services for employees are curated to foster accessibility and efficiency. Beyond facilitating external email access, services like that enable employees and beneficiaries to seamlessly engage with the eligibility and referral system and other health outreach services.

The 'Patient e-Services' platform, facilitated by the 'Careware application'. This application serves as a pivotal tool for both patients and the medical staff, centralizing and streamlining access to the health information system(HIS). However, as the hospital keeps looking into technological advancements in service delivery, it encounters inherent security challenges that must be addressed.

## **Healthcare infrastructure**

The Careware app is the most critical application to hospital functions. It is hosted in the hospital's data centre and relies mainly on two components: careware-app and careware-db. However, some supportive applications integrate with the careware solution to provide a unified console for hospital functions. The hospital infrastructure is designed based on multiple layers of technologies: Server-farm switches where servers are connected, Core switches, Edge gateway, Distribution switches, and Access switches that connect endpoint devices. Careware is accessed and supported by hospital infrastructure as illustrated in Figure 6.6. The infrastructure is designed to facilitate access to hospital digital services for medical devices, as well as staff workstations using access switches. These access switches are connected to multiple switches called 'Distribution switches' that are scattered over the hospital's buildings. The core layer consists of two modular switches that connect users to servers-farm switches, where hospital services are located, including careware applications. However, the core layer also provides access to the internet and WAN using an edge gateway(aka next-generation firewall). Additionally, the hospital is using a consolidated security deployment for the firewall, due to that the firewall provides two contexts, one to secure outbound traffic for the internet, while the second aims to control access to hosted services in the data centre.

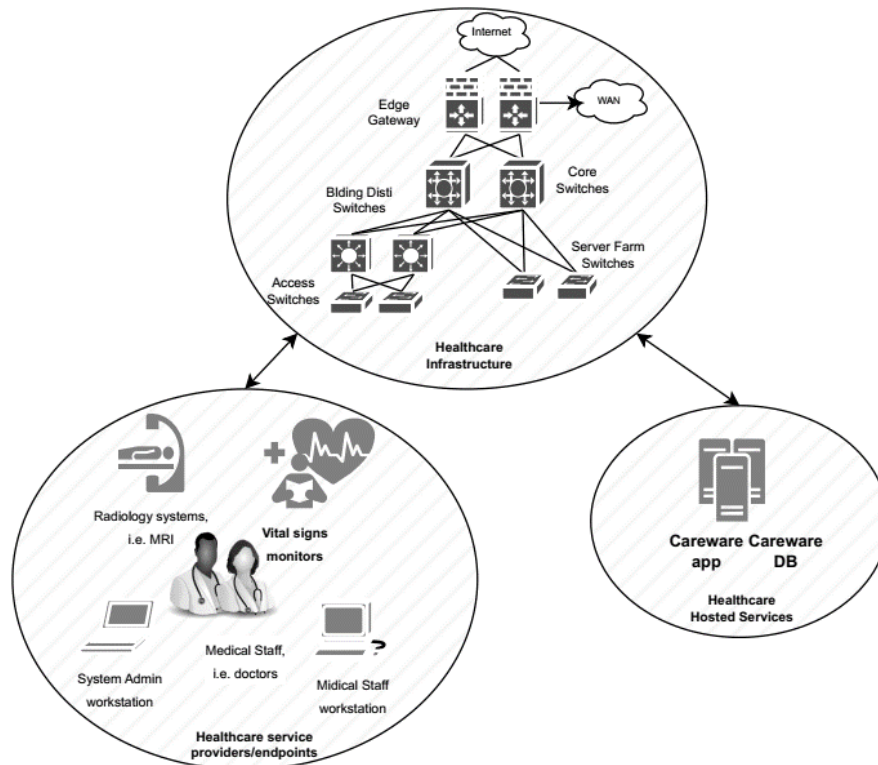


Figure 6.6. The infrastructure supporting the case study business.

## Healthcare challenges

Due to the continuous integration of modern technologies, the hospital has experienced an increase in cyberattacks over the past few months. There are two major types of these attacks. The first uses techniques like phishing and social engineering to get access to the Careware app by specifically targeting medical users and IT personnel. The second kind of attack uses a technique called service fingerprinting to determine the kind and version of software running on the hospital's network, which might leave that application vulnerable to attack. Considering these threats, the leadership of the hospital has decided to conduct a thorough security assessment with two distinct goals in mind: first, to expose any existing weaknesses or flows in the hospital's current digital infrastructure, particularly the Careware application; and second, to strengthen the hospital's cybersecurity assurance against future threats.

### 6.4.2. d-TM Implementation

The d-TM application in this case study is empowered by the d-TM platform. The head of medical services and two other hospital employees have been engaged to employ the d-TM concept. The following sections illustrate how the process unfolds and the data that is produced at each stage of using the d-TM model.

## Activity 1 Data Collection

This is the initial activity in the d-TM process; it aims to collect the required understanding of the case study digital services and supporting IT infrastructure. This information is collected with the support of the hospital’s systems team. Please note that only the systems configuration is intended to be gathered in this process, no patient records are needed or collected. This activity is focused on two processes:

- **Identifying Critical Services**

In the data collection activity, the initial step involves identifying crucial services. This step's objective is to understand the current business operations and the associated digital services. For this sake, information concerning operational services is gathered using the d-TM tool. To input data into this tool, we utilize the service collection form, which was elaborated upon in a previous chapter. Both the research team and the hospital system team handle data collection. Upon evaluation, it was determined that the patient e-service, known as Careware, stands out as the hospital's most vital asset in delivering services. As criticality of hospital services and their significance in service delivery are inputted into the d-TM tool. Figure 6.7 illustrates the results of this process. *As a result, subsequent evaluations will primarily focus on the Careware application.*

#	Service ID#	Service Name	Service Description	Service Criticality
37	B05	General e-services	The hospital's general e-services span across four critical categories. From a certificate verification system, ensuring the authenticity of patients' medical credentials, to a dedicated platform for health education, the hospital endeavours to disseminate critical healthcare knowledge.	Medium
39	B03	Patient e-services	A dedicated e-service portal serves as a digital gateway. Patients can effortlessly reschedule or cancel appointments, access comprehensive medical histories, retrieve previous laboratory test results, immunisation records, medication prescriptions, and much more. This 24/7 digital assistance proves invaluable, especially during emergencies, rendering services at the patient's convenience.	High
40	B04	Employee e-Services	Recognizing the pivotal role of its workforce, the hospital's e-services for employees are curated to foster accessibility and efficiency. Beyond facilitating external email access, services like PAC enable employees and beneficiaries to seamlessly engage with the eligibility and referral system and other health insurance services.	Medium
42	B06	Suppliers e-Services	Fostering transparency and efficacy, the hospital's e-services for suppliers streamline the procurement process. With pivotal services like Bids and Supplier, suppliers navigate a transparent bidding process, facilitated by the government's "Ginad" procurement initiative. This advanced grade scheme adheres to stringent guidelines, assuring an equitable and meticulous bidding environment.	Not Critical

Figure 6.7. The output table of identified services supporting the hospital business .

The Figure presents a ‘Business Services Table,’ which demonstrates collected services, each service is accompanied by a brief description that provides insight into its function, and their criticality to the hospital functions. Four distinct services are illustrated: General Services, Patient Services, Employee Services, and Supplier Services. These services satisfy various stakeholders in the hospital ecosystem – patients, employees, and suppliers. The services can be described as follows, ‘General Services’ provides foundational functionalities like the storage of patients' medical credentials and healthcare knowledge dissemination. While ‘Patient Services’ present the most comprehensive, offering a wide range of facilities from routine check-ups to emergency services. Moreover, ‘Employee Services’ focuses on making the healthcare system more accessible and efficient for hospital staff. Finally, ‘Supplier Services’ is involved with the procurement process and supplier interactions, hinting at transparency and efficiency.

Eventually, Patient Services' stands out as being of 'High' criticality. This is consistent with expectations, as services that directly impact patient health and well-being would logically be paramount. While both 'General Services' and 'Employee Services' are deemed 'Medium' criticality. While they play significant roles, they might not have the direct and immediate impact on patient health as 'Patient Services' does. Lastly, 'Supplier Services' is categorized as 'Not Critical'. This might suggest that while procurement processes are important, they do not directly or immediately impact patient care.

- **Identifying Surrounding Infrastructure**

Having collated information on business services in the preceding stage, it becomes imperative to determine the digital infrastructure supporting these critical services. Within the sequence of d-TM threat analysis, our subsequent action is to pinpoint the foundational infrastructure supporting the e-patient service called 'Careware'. Notably, Careware has been distinguished as being of paramount significance for the hospital's operational continuity. The methodology to analyse the infrastructure surrounding the Careware application is outlined in three sequential steps:

### **Step 1 Identifying asset details**

This step is dedicated to gathering detailed information about infrastructure. The information is collected by the d-TM tool. This includes information such as the asset's name, its brand, the software version, and the associated configuration file. Moreover, every asset is systematically aligned with the appropriate d-TM attack layer and is given a unique identification number to be identified in the next stages. The table depicted in Figure 6.8 embodies the outcomes of this stage. In total, thirteen assets have been identified, all of which underlie the Careware application.



Service Data Assets Table						
Below find assets list						
#	Service. ID(#)	Asset. ID(#)	Asset. Name	Asset. Brand	Asset. SW Version	Asset Config File
34	Patients e-services (Bs3)	Ag1	User_Browser01	MS Edge	112	collection/config/Agent_EndPoint_-_BUser01.txt
35	Patients e-services (Bs3)	Ag2	Opert_Browser01	Firefox	115	collection/config/Agent_EndPoint_-_Oper01.txt
36	Patients e-services (Bs3)	Net1	GF_Access_Switch01	Juniper EX-3400	Junos 12.3	collection/config/EHS-KFH-AS0-1-SW-01.txt
37	Patients e-services (Bs3)	Net2	BLD01_Disti_Switch01	Juniper EX-4650	Junos 13.3	collection/config/EHS-KFH-MB-GF-DS-0.txt
38	Patients e-services (Bs3)	Net3	Core_Switch01	Juniper EX-4650	Junos 13.3	collection/config/EHS-KFH-DC-CS-0.log
39	Patients e-services (Bs3)	Net4	Edge_Firewall_01	Juniper SRX-380	Junos 21.2	collection/config/EHS-KFH-DC-EFW-0.txt
40	Patients e-services (Bs3)	Net5	Server_Farm_SW01	Juniper EX-4650	Junos 18.1	collection/config/EHS-KFH-DC-SF-0.log
41	Patients e-services (Bs3)	Cmp1	CarewareAPP_Server01	DELL R750	Linux Centos 7	collection/config/Careware_App_Compute.txt
42	Patients e-services (Bs3)	Cmp2	CarewareDB_Server01	DELL R750	Linux Centos 7	collection/config/Careware_Db_Compute.txt
43	Patients e-services (Bs3)	App1	CarewareAPP	Apache	2.4	collection/config/Careware_App.txt
44	Patients e-services (Bs3)	App2	CarewareDB	Oracle	23.2	collection/config/Careware_DB.txt
45	Patients e-services (Bs3)	Stg1	CarewareDB_Server01	DELL R750	Linux Centos 7	collection/config/Careware_Stg_DB.txt
46	Patients e-services (Bs3)	Ag3	Terminal_Client	MobaXterm	22.3	collection/config/Agent_EndPoint_-_Oper01_Me8apx9.txt

Figure 6.8. The output table of infrastructure asset details supporting the patient e-service.

The Figure details a list of assets associated with a specific service, notably ‘patient e-service’. The assets cover a broad range, from browsers (e.g., User\_Browser01 and Opert\_Browser01) to network switches (e.g., GF\_Access\_Switch01), firewalls, servers, and applications. This proposes a multi-tiered infrastructure with a combination of user interfaces, networking equipment, and backend servers.

The table also captures a variety of brands, like MS Edge, Firefox, Juniper, DELL, Apache, and Oracle, reflecting a heterogeneous environment. Versions, such as the Juniper devices(e.g., junos 21.2). The presence of configuration file paths (e.g., collection/config/Agent\_EndPoint...) implies that configurations for each asset are imported and stored for easier retrieval. However, Each asset has been assigned a unique ID based on the d-TM concept, facilitating easier referencing and lookup. This is essential for efficient asset management and tracking.

## Step 2 Identifying asset administration details

The second step seeks to gather in-depth details about assets registered in the prior stage. By obtaining this supplementary information, we gain enhanced insights into the administrative properties of these assets. Such data is invaluable as it enables security analysts to recognize the mechanisms instituted for managing these resources. Figure 6.9 represents the outcome of this

step. This outcome enumerates the registered assets, exposing further details such as management IP addresses, ports, and the conditions required for access.

**Data Assets mgmt Table**  
Below find assets mgmt list

#	Asset. ID(#)	Mgmt. IP	Mgmt. Port	Mgmt. Agent	Mgmt. Access	Mgmt. Privilage
13	User_Browser01 (Ag1)	N/A	N/A	User_Browser01 (Ag1)	Direct_Access	User
14	Opert_Browser01 (Ag2)	N/A	N/A	Opert_Browser01 (Ag2)	Direct_Access	User
15	Terminal_Client (Ag3)	N/A	N/A	Terminal_Client (Ag3)	Direct_Access	User
16	GF_Access_Switch01 (Net1)	10.128.134.20	22	Terminal_Client (Ag3)	Local_Network	Admin
17	BLD01_Disti_Switch01 (Net2)	10.128.134.30	22	Terminal_Client (Ag3)	Local_Network	Admin
18	Core_Switch01 (Net3)	10.128.134.100	22	Terminal_Client (Ag3)	Local_Network	Admin
19	Edge_Firewall_01 (Net4)	10.128.134.220	22, 443, 23	Terminal_Client (Ag3)	Local_Network	Admin
20	Server_Farm_SW01 (Net5)	10.128.134.90	22	Terminal_Client (Ag3)	Local_Network	Admin
21	CarewareAPP_Server01 (Cmp1)	10.128.137.51	22	Terminal_Client (Ag3)	Local_Network	Admin
22	CarewareDB_Server01 (Cmp2)	10.128.137.57	22,443	Terminal_Client (Ag3)	Local_Network	Admin
23	CarewareAPP (App1)	10.128.137.51	443	Opert_Browser01 (Ag2)	Local_Network	Admin
24	CarewareDB (App2)	10.128.137.57	443	Opert_Browser01 (Ag2)	Local_Network	Admin
25	CarewareDB_Server01 (Stg1)	10.128.137.57	22	Terminal_Client (Ag3)	Direct_Access	Admin

Figure 6.9. The output table of asset administration details supporting the patient e-service.

The ‘Data Assets mgmt. Table’ offers a structured view into the management details of various assets. This table is evidently categorized to deliver essential information about each asset's administrative tasks. Each asset is labeled with an ID, name, and an IP address. This address is used to identify the asset in the network and facilitate asset accessibility and management. While some assets like ‘User\_Browser01’ and ‘Opert\_Browser01’ do not have IP addresses or ports specified (the reason is that assets are client-end tools cannot assign a dedicated IP).

Furthermore, Management Agent column details the software or entity that is primarily responsible for interacting with the asset. For instance, 'Terminal\_Client' is a tool used for accessing most of infrastructure assets for administrative tasks. Management Access column describes the type of access mechanism in place. Most assets either have 'Direct\_Access' or 'Local\_Network', indicating whether they are accessed directly or via a local network. Management Privilege column details the level of authority or access given to the entity managing the asset. Most assets are managed with ‘Admin’ privileges, which suggests they are crucial and have higher-level of controls. A few assets, especially those with ‘Direct\_Access’, are managed with ‘User’ privileges, possibly indicating more general or low-level tasks.

### Step 3 Asset Dependency

The final stage in the asset identification process delves into understanding asset dependencies. Recognizing the interdependencies of assets is pivotal for the d-TM process, as assets often need to interact with one another to provide access to digital resources. This interrelation, when comprehended thoroughly, allows the d-TM platform to generate data-flow diagrams effectively. However, the nature of these dependencies is often shaped by the type of asset and its specific access prerequisites. Figure 6.10 provides a representation of the outcomes of this stage. This figure explains the relationships between assets, indicating their dependencies and type of access. For example, the Careware application is reliant on two assets: ‘CarewareAPP\_Server01’ and ‘CarewareDB Application’. The former serves as the direct hosting platform, while the latter facilitates processing and is accessed remotely via the network.

#	Asset ID(#)	Dep. Asset	Dep. Type	Dep. Access
20	User_Browser01 (Ag1)	GF_Access_Switch01 (Net1)	Exchange	Local_Network
21	Opert_Browser01 (Ag2)	GF_Access_Switch01 (Net1)	Exchange	Local_Network
22	GF_Access_Switch01 (Net1)	User_Browser01 (Ag1)	Exchange	Local_Network
23	GF_Access_Switch01 (Net1)	Opert_Browser01 (Ag2)	Exchange	Local_Network
25	GF_Access_Switch01 (Net1)	BLD01_Disti_Switch01 (Net2)	Exchange	Local_Network
26	BLD01_Disti_Switch01 (Net2)	GF_Access_Switch01 (Net1)	Exchange	Local_Network
27	BLD01_Disti_Switch01 (Net2)	Core_Switch01 (Net3)	Exchange	Local_Network
28	Core_Switch01 (Net3)	BLD01_Disti_Switch01 (Net2)	Exchange	Local_Network
29	Core_Switch01 (Net3)	Edge_Firewall_01 (Net4)	Exchange	Local_Network
30	Core_Switch01 (Net3)	Server_Farm_SW01 (Net5)	Exchange	Local_Network
31	Edge_Firewall_01 (Net4)	Core_Switch01 (Net3)	Exchange	Local_Network
32	Edge_Firewall_01 (Net4)	Server_Farm_SW01 (Net5)	Exchange	Local_Network
33	Server_Farm_SW01 (Net5)	Core_Switch01 (Net3)	Exchange	Local_Network
34	Server_Farm_SW01 (Net5)	Edge_Firewall_01 (Net4)	Exchange	Local_Network
35	CarewareAPP_Server01 (Cmp1)	Server_Farm_SW01 (Net5)	Exchange	Local_Network
36	CarewareDB_Server01 (Cmp2)	Server_Farm_SW01 (Net5)	Exchange	Local_Network
37	CarewareAPP (App1)	CarewareAPP_Server01 (Cmp1)	Host	Direct_Access
39	CarewareAPP (App1)	CarewareDB (App2)	Process	Local_Network
40	CarewareDB (App2)	CarewareDB_Server01 (Cmp2)	Host	Direct_Access
41	CarewareDB (App2)	CarewareAPP (App1)	Process	Local_Network
42	CarewareDB (App2)	CarewareDB_Server01 (Sig1)	Storage	Direct_Access
43	CarewareDB_Server01 (Sig1)	CarewareDB_Server01 (Cmp2)	Host	Direct_Access
45	CarewareDB (App2)	CarewareAPP (App1)	Exchange	Local_Network
46	Terminal_Client (Ag3)	GF_Access_Switch01 (Net1)	Exchange	Local_Network
47	Server_Farm_SW01 (Net5)	CarewareAPP_Server01 (Cmp1)	Exchange	Local_Network
48	Server_Farm_SW01 (Net5)	CarewareDB_Server01 (Cmp2)	Exchange	Local_Network
49	GF_Access_Switch01 (Net1)	Terminal_Client (Ag3)	Exchange	Local_Network

Figure 6.10. The output table of asset dep. details supporting the patient e-service.

The figure represents ‘Data Assets Dependency Table’, which offers a comprehensive overview of how different assets within a system are interdependent. This interconnection is crucial for understanding the flow and exchange of data, especially in complex digital environments. Each entry in the table is associated with an identification number. This identifier denotes a specific asset within the system. Three key columns follow the Asset ID:

- **Dep. Asset:** This column specifies which assets the primary asset is dependent upon. In simpler terms, it indicates the assets that the primary one interacts with or relies upon for its functionality.

- **Dep. Type:** This column categorizes the nature of dependency. Most dependencies in the table are labelled 'Exchange', suggesting a two-way communication or interaction between the assets. It represents assets that work in sharing or exchanging data to accomplish tasks.
- **Dep. Access:** This column defines how the primary asset accesses its dependent asset. For instance, 'Local Network' implies that the asset is accessed over a shared internal network, while 'Direct Access' indicates a more immediate connection.

For instance, the asset with ID 'User\_Browser01 (Ag1)' depends on the asset 'GF\_Access\_Switch01 (Net1)' through an 'Exchange' type of dependency, and this access happens over a 'Local Network'. Technically, the end-user Ag1(web browser) exchanges data with a network switch that shares these data to other network elements till it reaches the destination digital resource.

## **Activity 2 Data Analysis**

The Data analysis is the second activity in the d-TM process. This activity aims to analyse collected data from the data collection activity to identify data levels and phases for each asset. Furthermore, during this activity, the d-TM platform is employed to produce a data flow diagram. This diagram maps out the specific data assets that support the patient e-service, capturing intricate details and interdependencies. This activity is outlined into two processes:

- **Identify Data Level and Phases**

This process is focuses on data and aims to identify and classify collected asset data to the d-TM levels and phases which will be analysed in the next stages. Each asset is evaluated based on the three d-TM data levels concept: management, Control, and Business. Once data levels are identified for each asset, these identified data levels are inspected to determine the appropriate data phase. Data phases are categorized as data at-rest, in-process, and in-transit. The process is relying on assessing collected asset config files. Figure 6.11 represents a sample output of the produced result, which is due to the enormous size of the output table of this activity. The outcome of this process reveals forty-three entries representing assets to data-level/phase mapping. The data level to phases analysis shows that management-data as most targeted data for attention, then Business, and lastly Control data. Management data level is crucial to be secured due to the impact could cause if it got compromised. In addition to the identified data-levels, data phases are analysed as a part of the process. It is important to be identified due to the valuable insights that could bring later to the mitigation process.

**Identified Data Level and phases Table**  
 Below find assets, config and its related data levels and phases

Asset ID	Data Level	Data Phase
⌵ Ag1	mD	Dr
⌵ Ag1	bD	Dr
<i>Asset Config : - auto fill is enabled with app credentials</i>		
⌵ Ag3	mD	Dr
<i>Asset Config : - remote access software with auto login to multiple workstations</i>		
⌵ Ag3	cD	Dp
<i>Asset Config : - not uptodate browser used to access web managment network devices consoles</i>		
⌵ Net1	mD	Dt
<i>Asset Config : set system services ftp</i>		
⌵ Net1	mD	Dt
<i>Asset Config : set system services telnet connection-limit 5</i>		
⌵ Net1	mD	Dr
⌵ Net1	bD	Dt
⌵ Net2	mD	Dr
⌵ Net2	mD	Dt
⌵ Net1	mD	Dt
⌵ Net2	bD	Dt
⌵ Net2	cD	Dp
⌵ Net3	mD	Dr

Figure 6.11. Sample output table of assets data levels and phases for the patient e-service.

The figure presents a structured overview of data levels and phases associated with various assets. The table consists of three primary columns: Asset ID, Data Level, and Data Phase. As well as associated configurations denoted as ‘Asset Config’ in each entry. These configurations provide more detailed insights into the concerning data of the assets. For instance: ‘set system services ftp’ and ‘set system services telnet connection-limit 5’ are configurations specified for ‘Net1’ and

'Net2' respectively. In addition to the Asset identification, Data levels and phases of each asset are presented in second and third columns as follows:

**Data Level:** The data level provides context to the type of data related to each asset. The classifications and abbreviations are as follow:

- mD: Management Data, which refers to data concerning the asset's own administration and operation. This could encompass access configuration, settings, and operational parameters.
- cD: Control Data, which compresses data exchanges occurring between systems. This can provide insights into the interdependencies and interactions between different assets or systems.
- bD: Business Data, representing data originating from business end-users and directed towards business services. This may include transactional data, user inputs, or service requests.

**Data Phase:** It specifies the state in which the data exists. The possible phases include:

- Dr: Data at-rest, indicating the data is stationary and not being actively used or transferred.
- Dp: Data in-process, signifying data is currently being processed or used.
- Dt: Data in-transit, denoting data is being transferred from one location to another.

- **Construct Data-flow Diagram**

The d-TM platform incorporates a pivotal feature that allows users to automatically generate Data Flow Diagrams (DFDs) for infrastructures supporting critical services. This capability draws upon the accumulated data about assets and their interdependencies to construct these DFDs. Figure 6.12 demonstrates the output of this process for patients e-service, showcasing interconnected assets and their relational dependencies. Each dependency is denoted atop the connecting links between assets. For example, 'Ag1' has a data exchange (denoted as 'E') relationship with 'Net1', with the flow direction also indicated. Further, the DFD is structured into clusters based on the d-TM attack layers. Each cluster is distinguished by a unique colour code. To illustrate, the 'Agent' layer is represented in a 'Gray' hue. The diagram's narrative begins with a 'USER' who utilizes the 'Agent' software to access 'Network' resources. These resources interface with 'Compute' entities that host 'Application' components, which in turn, manage data within the 'Storage' domain.

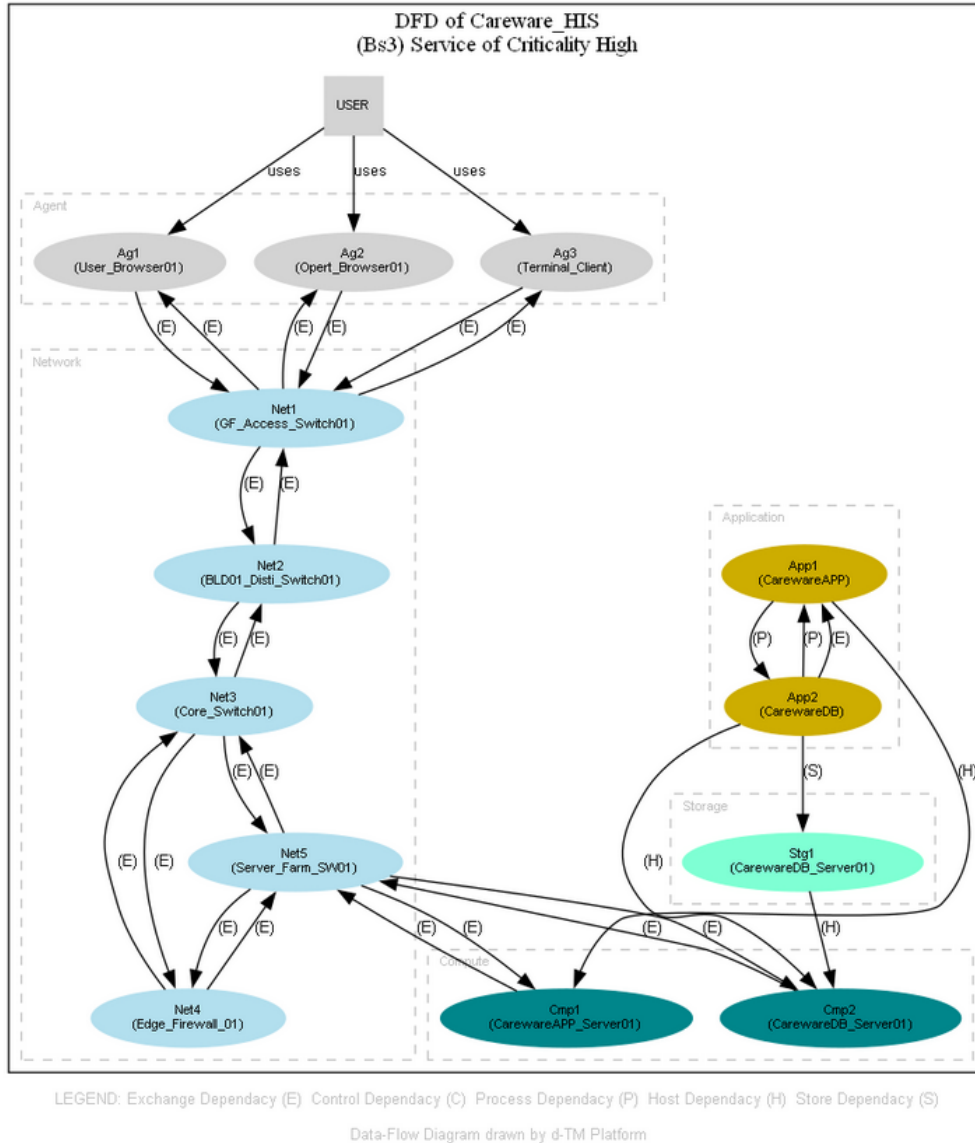


Figure 6.12. The output DFD of the Patient e-service.

The DFD is detailed and offers an insightful overview of the patient e-service infrastructure. The hierarchical structure, starting from the user and delving into storage, provides a clear sequence of interactions and dependencies. However, the DFD is a coherent interpretation of how various components of an IT system interact and depend on one another. The DFD blocks can be described as follows:

- User Interaction and Agent Layer:

At the topmost layer, the diagram displays an actor labelled ‘USER.’ This user represents medical staff or a system administrator. Notably, there are three potential agent nodes (Ag1, Ag2, and Ag3) that user interacts with. It is visible from the accompanying information that while Ag1 is utilized by medical staff to access the Patient e-service, both

Ag2 and Ag3 are used by system administrators for operational tasks. This divergence is vital as it determines the functions and access privileges of different user roles.

- Network Layer:

Descending the diagram, the network layer consists of five primary assets (Net1 to Net5). These are critical assets that facilitate data exchange, ensuring seamless connectivity to the subsequent compute layer. The inclusion of elements like switches (GF\_Access\_Switch01, BLD01\_Disti\_Switch01, etc.) and firewalls (Edge\_Firewall\_01) indicates a multi-tiered, possibly segmented network design – a best practice in ensuring network security and optimizing traffic.

- Application & Compute Layer:

There are two distinct nodes: App1 (CarewareAPP) and App2 (CarewareDB), signifying the application and its associated database, respectively. The Careware application gives its interconnections to multiple preceding layers. Two compute servers, Cmp1 and Cmp2, host the application and database. It is evident from the connections that the database server (Cmp2) is particularly crucial, given it is connected to both the application and storage layers.

- Storage Layer:

The last layer indicates the storage node (Stg1). This storage is directly linked to the Cmp2 server, presenting local storage rather than network-attached storage (NAS) or storage area network (SAN). This could imply faster data retrieval times but could also be a potential single point of failure if not managed correctly.

The DFD provides a concise legend detailing diverse types of dependencies, from data exchanges (E) to control, process, host, and storage dependencies. These legends are vital for a comprehensive understanding of data flow, potential bottlenecks, and security flaws.

### **Activity 3 Threat Analysis**

The threat analysis within the d-TM process is a crucial phase, aiming to examine the information gathered in earlier activities to recognize weaknesses, potential threats, and the subsequent impact they may have on business continuity. Given the complicated nature of digital ecosystems, such a systematic approach is imperative to ensure robust threat management. The three sequential steps of the threat analysis are explained as follows:

#### **Step 1 Identify Weaknesses**

This is the first step in the threat analysis activity. This stage is intended to examine collected data to identify potential weaknesses. Leveraging detailed insights from identified data levels, phases, and associated configurations, the platform identifies a variety of weaknesses that can pose tangible risks to the operational continuity of the business. As depicted in Figure 6.13, a thorough



analysis revealed a total of forty-six potential weaknesses embedded within the data asset, each of which carries implications for business continuity.

A deeper dive into these weaknesses reveals a skewed distribution across different data levels. The Management-data level emerges as the most vulnerable, showing thirty-five out of the forty-six identified weaknesses. This suggests that administrative practices and operations might be particularly exposed. On the other hand, the Control-data level, representing system interactions, displayed two weaknesses. Meanwhile, the Business-data level, which pertains to end-user operations and business services, manifested nine specific weaknesses. This distribution underscores the need for a heightened focus on the Management-data level, while also considering the weaknesses present in the other data levels.

Filter Clear Exit Export to PDF

Identified Data Level Weaknesses Table  
Below find assets, config and its related data levels weaknesses

Asset ID	Data Level	Data Phase	Weakness Name
∨ Ag1	mD	Dr	Improper Access Control
∨ Ag1	bD	Dr	Use of Web Browser Cache Containing Sensitive Information
∨ Ag1	mD	Dr	Weak Password Requirements
∨ Ag1	bD	Dp	Insecure Storage of Sensitive Information
∨ Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control
∨ Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor
<b>Asset Config</b> : - remote access software with auto login to multiple workstations			
<b>Weakness ID</b> : 200			
<b>Weakness Details</b> : The product exposes sensitive information to an actor that is not explicitly authorized to have access to that information.			
∨ Ag3	cD	Dp	Dependency on Vulnerable Third-Party Component
∨ Ag1	mD	Dr	Use of Weak Credentials
∨ Ag2	bD	Dr	Insecure Storage of Sensitive Information
∨ Ag2	cD	Dp	Dependency on Vulnerable Third-Party Component
∨ Net1	mD	Dt	Cleartext Transmission of Sensitive Information
∨ Net1	mD	Dr	Reliance on Insufficiently Trustworthy Component
<b>Asset Config</b> : set system services ftp			
<b>Weakness ID</b> : 1357			
<b>Weakness Details</b> : The product is built from multiple separate components, but it uses a component that is not sufficiently trusted to meet expectations for security, reliability, updateability, and maintainability.			
∨ Net1	mD	Dt	Cleartext Transmission of Sensitive Information
∨ Net1	mD	Dt	Reliance on Insufficiently Trustworthy Component
∨ Net2	bD	Dt	Improper Access Control
∨ Net2	mD	Dr	Reliance on Insufficiently Trustworthy Component
∨ Net2	mD	Dt	Cleartext Transmission of Sensitive Information
∨ Net3	mD	Dr	Reliance on Insufficiently Trustworthy Component
∨ Net3	mD	Dt	Cleartext Transmission of Sensitive Information
∨ Net3	mD	Dt	Reliance on Insufficiently Trustworthy Component
∨ Net3	mD	Dt	Cleartext Transmission of Sensitive Information
∨ Net4	mD	Dr	Reliance on Insufficiently Trustworthy Component
∨ Net4	mD	Dt	Cleartext Transmission of Sensitive Information
∨ Net4	mD	Dp	Execution with Unnecessary Privileges
∨ Net4	mD	Dt	Cleartext Transmission of Sensitive Information
∨ Net4	mD	Dt	Use of a Broken or Risky Cryptographic Algorithm

Figure 6.13. Sample output table of identified weaknesses exists in the patient e-service.

The figure represents a table of identified weaknesses associated with specific assets within patient e-service system, which provides invaluable insights for IT security professionals. The table is structured with distinct columns - Asset ID, Data Level, Data Phase, and Weakness Name. This layout suggests a thorough effort to categorize weaknesses based on assets, the kind of data they handle, the phase of data (whether at-rest, in-process, or in-transit), and the specific weakness pertinent to the asset. Apparently, A single asset can have multiple weaknesses. For instance, the asset 'Ag1' is associated with different weaknesses across various data levels and phases. This highlights the complicated nature of threat landscapes where an asset might be susceptible to multiple threats across its lifecycle. On the other side, there are recurring weaknesses across different assets, such as 'Insecure Storage of Sensitive Information' and 'Improper Access Control.' This could indicate systemic issues or common weak configurations present in multiple components of the system. Furthermore, the analysis does not just list generic weaknesses but provides specific issues such as 'Weak Password Requirements' or 'Exposure of Sensitive Information to an Unauthorized Actor'.

In summary, the displayed table indicates a pronounced vulnerability within the Management-data (mD) level, highlighting potential risks in the administrative and operational facets of the system. These could be primary areas of interest for malicious entities. Additionally, the data phase underscores the inherent risks associated with transferring sensitive information through channels susceptible to unauthorized interception.

## **Step 2 Identify Threats**

This step involves the identification of threats stemming from weaknesses determined in the prior step. Each detected weakness is evaluated to identify the associated threat to the data asset. As a result of this evaluation, forty potential threats have been identified due to the presence of these weaknesses. Predominantly, these threats target the Management-data level, accounting for 32 out of the 40 threats, while other threats are distributed among Control and Business data-level. Due to the table size, Figure 6.14 presents a subset of these threats with accompanying specifics. Important to mention, that the discrepancy between the number of identified weaknesses and the subsequent number of threats can arise due to some weaknesses might be repeated across different assets, data levels, or data phases, but they all lead to the same type of threat. So, these weaknesses are ignored intentionally as the aim is to demonstrate the model applicability and avoiding overwhelming results.

Filter by Asset ID:

Filter Clear Exit Export to PDF

Identified Data Level Threats Table  
Below find assets, config, weakness and its related data levels threats

Asset ID	Data Level	Data Phase	Weakness Name	Threat Name
⌘ Ag1	mD	Dr	Improper Access Control	Malicious Logic Insertion
⌘ Ag1	bD	Dr	Use of Web Browser Cache Containing Sensitive Information	Retrieve Embedded Sensitive Data
⌘ Ag1	mD	Dr	Improper Access Control	Modification of Windows Service Configuration
⌘ Ag1	mD	Dr	Weak Password Requirements	Password Brute Forcing
⌘ Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data
⌘ Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Network Boundary Bridging
⌘ Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor	Retrieve Embedded Sensitive Data
⌘ Ag3	cD	Dp	Dependency on Vulnerable Third-Party Component	Client-side Injection-induced Buffer Overflow
⌘ Ag2	bD	Dr	Insecure Storage of Sensitive Information	Retrieve Embedded Sensitive Data
⌘ Ag3	cD	Dp	Dependency on Vulnerable Third-Party Component	Malicious Logic Insertion
⌘ Net1	mD	Dt	Cleartext Transmission of Sensitive Information	Interception

**Asset Config :** set system services ftp

**Weakness ID :** 319  
**Weakness Details :** The product transmits sensitive or security-critical data in cleartext in a communication channel that can be sniffed by unauthorized actors.

**Threat ID :** 117  
**Threat Details :** An adversary monitors data streams to or from the target for information gathering purposes. This attack may be undertaken to solely gather sensitive information or to support a further attack against the target. This attack pattern can involve sniffing network traffic as well as other types of data streams (e.g. radio). The adversary can attempt to initiate the establishment of a data stream or passively observe the communications as they unfold. In all variants of this attack, the adversary is not the intended recipient of the data stream. In contrast to other means of gathering information (e.g., targeting data leaks), the adversary must actively position themselves so as to observe explicit data channels (e.g. network traffic) and read the content. However, this attack differs from a Adversary-In-the-Middle (CAPEC-94) attack, as the adversary does not alter the content of the communications nor forward data to the intended recipient.

⌘ Net1	mD	Dr	Reliance on Insufficiently Trustworthy Component	Password Brute Forcing
--------	----	----	--	------------------------

Figure 6.14. Sample output table of identified threats targeting the Careware system.

The above figure represents a table that links identified data levels and phases to their respective weaknesses and associated threats. Where each asset is clearly identified by its unique Asset ID, which aids in tracking and managing weaknesses and threats specific to each asset. The weaknesses identified are explicitly mapped to potential threats. This mapping offers a direct relationship between weaknesses in the system and potential exploitations. For example, a weakness like ‘Improper Access Control’ can lead to a threat like ‘Malicious Logic Insertion’. Some weaknesses appear multiple times across different assets, indicating systemic issues. For instance, the ‘Retrieve Embedded Sensitive Data’ weakness is recurrent, suggesting a pervasive issue in the protection of sensitive data across different assets. Furthermore, the table provides in-depth details for specific threats. For instance, Threat ID: 117 explains an adversary's potential actions and intentions, offering context to understand the threat's significance and mode of operation.

As a sum up of the Data Levels and Phases with their corresponding threats, below is a breakdown of the number of threats associated with each data level and phase, as well as some of the specific threats targeting them:

- **Management Data (mD):**
  - Data at Rest (Dr):
    - Improper Access Control
    - Cleartext Transmission of Sensitive Information
    - Reliance on Insufficiently Trustworthy Component
  - Data in Process (Dp):
    - Use of Weak Password Requirements
    - Exposure of Sensitive Information to an Unauthorized Actor
    - Execution with Unnecessary Privileges

- Data in Transit (Dt):
  - Interception
  - Application Fingerprinting
- **Business Data (bD):**
  - Data at Rest (Dr):
    - Improper Access Control
    - Missing Encryption of Sensitive Data
  - Data in Process (Dp):
    - Use of Password System for Primary Authentication
    - Improper Isolation or Compartmentalization
  - Data in Transit (Dt):
    - Browser-in-the-Middle (BITM)
- **Control Data (cD):**
  - Data in Process (Dp):
    - Dependency on Vulnerable Third-Party Component

### **Step 3** *Determine Threat Criticality*

Threat criticality is determined at this stage; this is the last step in the threat analysis process. The threats are evaluated seeking to understand their potential consequences on business continuity. This determination is facilitated by the d-TM evaluation criteria, which draws upon specific prerequisites for each threat. Subsequently, the d-TM tool conducts an assessment to discover the definitive threat criticality level. Figure 6.15 shows a representation of these assessments using a color-coded scheme, for clarity. From the prior analysis step, forty threats were detected, and their criticalities determined. Among these threats eleven are marked as ‘Dark Red ‘ indicating a ‘ potential impact on operations. Additionally, three are labeled as ‘Red ‘ signifying a ‘ impact. The remaining twenty-one threats are represented in shades of ‘Orange’. Yellow ‘ indicating impacts ranging from ‘Medium’, to ‘Very Low’, respectively.

Identified Data Level Threats Criticalities					
Below find assets, config and its related data levels threats Criticalities					
Asset ID	Data Level	Data Phase	Weakness Name	Threat Name	Threat Criticality
Ag1	mD	Dr	Improper Access Control	Malicious Logic Insertion	Very Low
Ag1	bD	Dr	Use of Web Browser Cache Containing Sensitive Information	Retrieve Embedded Sensitive Data	Medium
Ag1	mD	Dr	Improper Access Control	Modification of Windows Service Configuration	Very Low
Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High
Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Network Boundary Bridging	Very High
Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor	Retrieve Embedded Sensitive Data	Very High
Ag3	cD	Dp	Dependency on Vulnerable Third-Party Component	Client-side Injection-induced Buffer Overflow	Low
Ag2	bD	Dr	Insecure Storage of Sensitive Information	Retrieve Embedded Sensitive Data	Medium
Ag3	cD	Dp	Dependency on Vulnerable Third-Party Component	Malicious Logic Insertion	Very High
Net1	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very Low
Net1	mD	Dr	Reliance on Insufficiently Trustworthy Component	Password Brute Forcing	Low
Net1	mD	Dr	Reliance on Insufficiently Trustworthy Component	Application Fingerprinting	Very Low
Net1	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very Low
Net2	mD	Dr	Reliance on Insufficiently Trustworthy Component	Application Fingerprinting	Very Low
Net2	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very Low
Net3	mD	Dr	Reliance on Insufficiently Trustworthy Component	Application Fingerprinting	Very Low
Net3	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very Low
Net3	mD	Dt	Reliance on Insufficiently Trustworthy Component	Interception	Very Low
Net4	mD	Dr	Reliance on Insufficiently Trustworthy Component	Application Fingerprinting	Very Low
Net4	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very Low
Net4	mD	Dp	Execution with Unnecessary Privileges	Password Brute Forcing	Very High
Net4	mD	Dt	Use of a Broken or Risky Cryptographic Algorithm	Exploiting Incorrectly Configured SSL/TLS	Very High
Net4	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very High
Net4	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very High
Net4	mD	Dp	Improper Access Control	Exploiting Incorrectly Configured Access Control Security Levels	Very High
Net5	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very Low
Net5	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Low
Cmp1	mD	Dp	Improper Access Control	Application Fingerprinting	High
Cmp1	mD	Dp	Improper Access Control	Accessing Functionality Not Properly Constrained by ACLs	High
Cmp2	mD	Dp	Use of Password System for Primary Authentication	Brute Force	Medium
Cmp1	bD	Dt	Improper Isolation or Compartmentalization	Network Topology Mapping	Medium
Cmp2	mD	Dp	Use of Password System for Primary Authentication	Brute Force	Medium
App1	mD	Dp	Use of Single-factor Authentication	Authentication Abuse	Very High
App2	mD	Dp	Use of Password System for Primary Authentication	Authentication Bypass	High
App1	mD	Dp	Improper Access Control	Accessing Functionality Not Properly Constrained by ACLs	Very High

Figure 6.15. Sample output table of threat criticalities targeting Careware system.

The figure above presents a comprehensive analysis of identified threats across various data levels and phases, thoroughly linking them to specific weaknesses and subsequently measuring their threat criticality. Here below is a focused analysis of the produced critical threats for ‘Patient e-service’:

- **Top Critical Threats:**

The most recurring threats classified with 'Very High' criticality are 'Malicious Logic Insertion', 'Interception', 'Password Brute Forcing', and 'Retrieval of Embedded Sensitive Data'. Their presence suggests that the system is susceptible to both external attacks aiming to infiltrate or disrupt its operations, as well as to internal vulnerabilities that might expose sensitive information.

- **Most Critical Assets:**

The assets 'Ag3', 'Net4', and 'App1' stand out with multiple 'Very High' criticality threats. While threats to 'App2' and 'Cmp1' are critically 'high'. This indicates that these assets possess a severe risk that needs to be addressed.

- **Data Levels with the Most Threats:**

The 'mD' (Management Data) level consistently emerges with the highest frequency of 'Very High' criticality threats. This underscores the notion that administrative or operational aspects of the system might be its weak spot, serving as valuable targets for potential attackers.

- **Data Phases under Threat:**

- The 'Dr' (Data at-Rest) phase is recurrently associated with 'Very High' criticality threats, especially under the 'mD' data level. This suggests that stored data—whether it be configuration details, user information, or other vital data might be at significant risk.
- The 'Dt' (Data in-Transit) phase also gathers attention, especially with threats like 'Interception', implying vulnerabilities during data transfer.

In summation, it is evident that these dominant threats primarily target the Management-data level, specifically when data is in-Transit or at-Rest. While single threat aims to compromise Control-data level when in use.

## **Activity 4 Threat Mitigation**

In the d-TM threat analysis process, the concluding activity is threat mitigation. Its primary objective is to identify suitable controls for addressing critical threats. Moreover, this activity offers an assurance feature for controls, grounded in the d-TM criteria. This activity consists of two sequential steps:

### **Step 1 Determine suitable controls**

In the threat mitigation activity of the d-TM process, the initial step is the determination of appropriate controls. This stage aims to identify appropriate controls to address critical identified threats. The selection of these controls is essentially linked to the weakness's nature and high-priority threats determined in the preceding activities. The result of this evaluative stage is the recommendation of one or multiple controls uniquely adapted for each detected critical threat. The primary emphasis of this assessment is directed towards the top fourteen threats based on their severity and potential impact. Figure 6.16 presents the selected mitigation strategies aligned with each significant threat. Throughout this evaluative process, an aggregate of twenty-five controls was determined, and expressly suggested to protect data against threats deemed to have either "Very High" or "High" consequences. It is worth mentioning that the total number of controls is not mandatory to mirror the total number of threats, whereas multiple controls might be needed to adequately address a singular threat.

Determined Threat Controls						
Below find assets, config and its related data levels threats and Criticalities						
Asset ID	Data Level	Data Phase	Weakness Name	Threat Name	Threat Criticality	Threat Control
v Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High	Policy and Procedures
v Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High	Account Management
v Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High	Policy and Procedures
v Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Network Boundary Bridging	Very High	Policy and Procedures
v Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High	Account Management
v Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Network Boundary Bridging	Very High	Policy and Procedures
v Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor	Retrieve Embedded Sensitive Data	Very High	Policy and Procedures
v Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor	Retrieve Embedded Sensitive Data	Very High	Account Management
v Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor	Retrieve Embedded Sensitive Data	Very High	Policy and Procedures
v Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor	Retrieve Embedded Sensitive Data	Very High	Use-installed Software
v Ag3	iD	Dp	Dependency on Vulnerable Third-Party Component	Malicious Logic Insertion	Very High	Policy and Procedures
v Net4	mD	Dp	Execution with Unnecessary Privileges	Password Brute Forcing	Very High	Identification and Authentication (organizational Users)
v Net4	mD	Dp	Execution with Unnecessary Privileges	Password Brute Forcing	Very High	Remote Access
v Net4	mD	Dt	Use of a Broken or Risky Cryptographic Algorithm	Exploiting Incorrectly Configured SSL/TLS	Very High	Transmission Confidentiality and Integrity
v Net4	mD	Dt	ClearText Transmission of Sensitive Information	Interception	Very High	Remote Access
v Net4	mD	Dt	ClearText Transmission of Sensitive Information	Interception	Very High	Transmission Confidentiality and integrity
v Net4	mD	Dt	ClearText Transmission of Sensitive Information	Interception	Very High	Remote Access
v Net4	mD	Dt	ClearText Transmission of Sensitive Information	Interception	Very High	Transmission Confidentiality and Integrity
v Net4	mD	Dp	Improper Access Control	Exploiting Incorrectly Configured Access Control Security Levels	Very High	Remote Access
v Cmp1	mD	Dp	Improper Access Control	Application Fingerprinting	High	Access Enforcement   Restricted Access to Privileged Functions
v App1	mD	Dp	Use of Single-Factor Authentication	Authentication Abuse	Very High	Identification and Authentication (organizational Users)   Multi-Factor Authentication to Priv
v App1	mD	Dp	Use of Single-Factor Authentication	Authentication Abuse	Very High	Identification and Authentication (organizational Users)   Multi-Factor Authentication to Non
v App1	mD	Dp	Improper Access Control	Accessing Functionality Not Properly Constrained by ACLs	Very High	Access Enforcement
v App2	mD	Dp	Use of Password System for Primary Authentication	Authentication Bypass	High	Authenticator Management   Password Managers
v App2	mD	Dp	Use of Password System for Primary Authentication	Authentication Bypass	High	Reference Monitor

Figure 6.16. The output table of threat controls for the careware system.

The figure above provided a comprehensive details of threat controls associated with specific assets, data levels, and data phases. Assets like 'Ag3', 'Net4', 'Cmp1', 'App1', and 'App2' emerge as the primary subjects of focus for controls. The threats are subsequently ranked by their criticality, with many tagged as 'Very High', implying significant potential impacts. To address these threats, specific controls, such as 'Policy and Procedures', 'Remote Access', and 'Account Management', have been recommended. Some threats warrant multiple controls, demonstrating the intricate nature of the security solutions required. For instance, threats due to 'Clear-text Transmission of Sensitive Information' have controls emphasizing both 'Remote Access' and ensuring 'Transmission Confidentiality and Integrity'. In essence, this table serves as a strategic roadmap for organizations to identify, prioritize, and address threats with appropriate controls, ensuring that key assets remain secure and functional.

## Step 2 Determine control assurance level

In the d-TM threat mitigation activity, the last step encompasses the evaluation of control assurance levels. In this step, every previously identified control undergoes a thorough assessment based on the d-TM criteria. The d-TM criteria focus on three factors (completeness, complexity, and effectiveness). However, the objective is to determine the reliability and robustness of each control in addressing the associated threats.

The findings from the control assurance evaluation are visually represented in Figure 6.17. This figure details a color-coded table that designates the assurance level for every control. From the evaluation of nineteen controls, twelve emerged with a high assurance, depicted in the color 'Green'. On the contrary, seven controls were designated a moderate assurance level and are illustrated in 'Orange'.

Filter by Asset ID: Filter Clear Exit Export to PDF

Identified Threats Control Assurance Level  
Below find assets, config and its related data levels threats controls assurance

Asset ID	Data Level	Data Phase	Weakness Name	Threat Name	Threat Criticality	Threat Control	Control Assurance
Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High	Policy and Procedures	Moderate
Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High	Account Management	Moderate
Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Retrieve Embedded Sensitive Data	Very High	Policy and Procedures	Moderate
Ag3	mD	Dr	Storage of Sensitive Data in a Mechanism without Access Control	Network Boundary Bridging	Very High	Policy and Procedures	Moderate
Ag3	mD	Dr	Exposure of Sensitive Information to an Unauthorized Actor	Retrieve Embedded Sensitive Data	Very High	User-installed Software	High
Ag3	iD	Dp	Dependency on Vulnerable Third-Party Component	Malicious Logic Insertion	Very High	Policy and Procedures	High
Hes4	mD	Dp	Execution with Unnecessary Privileges	Password Brute Forcing	Very High	Identification and Authentication (organizational Users)	High
Hes4	mD	Dp	Execution with Unnecessary Privileges	Password Brute Forcing	Very High	Remote Access	High
Hes4	mD	Dt	Use of a Broken or Risky Cryptographic Algorithm	Exploiting Incorrectly Configured SSL/TLS	Very High	Transmission Confidentiality and Integrity	High
Hes4	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very High	Remote Access	High
Hes4	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very High	Transmission Confidentiality and Integrity	Moderate
Hes4	mD	Dt	Cleartext Transmission of Sensitive Information	Interception	Very High	Transmission Confidentiality and Integrity	Moderate
Hes4	mD	Dp	Improper Access Control	Exploiting Incorrectly Configured Access Control Security Levels	Very High	Remote Access	High
Cmp1	mD	Dp	Improper Access Control	Application Fingerprinting	High	Access Enforcement   Restricted Access to Privileged Functions	High
App1	mD	Dp	Use of Single-factor Authentication	Authentication Abuse	Very High	Identification and Authentication (organizational Users)   Multi-factor Authentication to Privileged	High
App1	mD	Dp	Use of Single-factor Authentication	Authentication Abuse	Very High	Identification and Authentication (organizational Users)   Multi-factor Authentication to Non-privile	High
App1	mD	Dp	Improper Access Control	Accessing Functionality Not Properly Constrained by ACLs	Very High	Access Enforcement	High
App2	mD	Dp	Use of Password System for Primary Authentication	Authentication Bypass	High	Authenticator Management   Password Managers	High
App2	mD	Dp	Use of Password System for Primary Authentication	Authentication Bypass	High	Reference Monitor	Moderate

Figure 6.17. The output table of control assurance levels for the careware system.

The figure above reveals that most threats are countered with "High" or "Moderate" control assurance. This indicates a relatively confident stance in the controls' efficacy for these threats. While "Moderate" assurance is not necessarily indicative of inefficacy, it is crucial to periodically review these controls, especially if they are guarding against "Very High" criticality threats. Moreover, "Remote Access Control", "Identification and Authentication (Organizational Users)", and "Access Enforcement" are among the controls frequently rated with "High" assurance.

Finally, the d-TM platform is empowered with multiple additional features that are utilized in this case study scenario as follows:

### Feature 1 d-TM Threat Analysis Dashboard

The d-TM dashboard provides an insightful visual presentation of various components related to threat analysis within the d-TM platform. Figure 6.18. Represent a dashboard that visually concludes the analysis of one critical service called 'Patient e-service'. The data is presented using a column chart format, with each bar denoting a specific detail associated with threats analysis process and their associated components. Also, the Figure shows two pie charts that provides a breakdown of the distribution of threats based on their assessed criticality. The categories range from "Very High" to "Very Low". The second one offers insights into the assurance level to manage threats through determined controls. The assurance level is categorized as High, Moderate, and low.



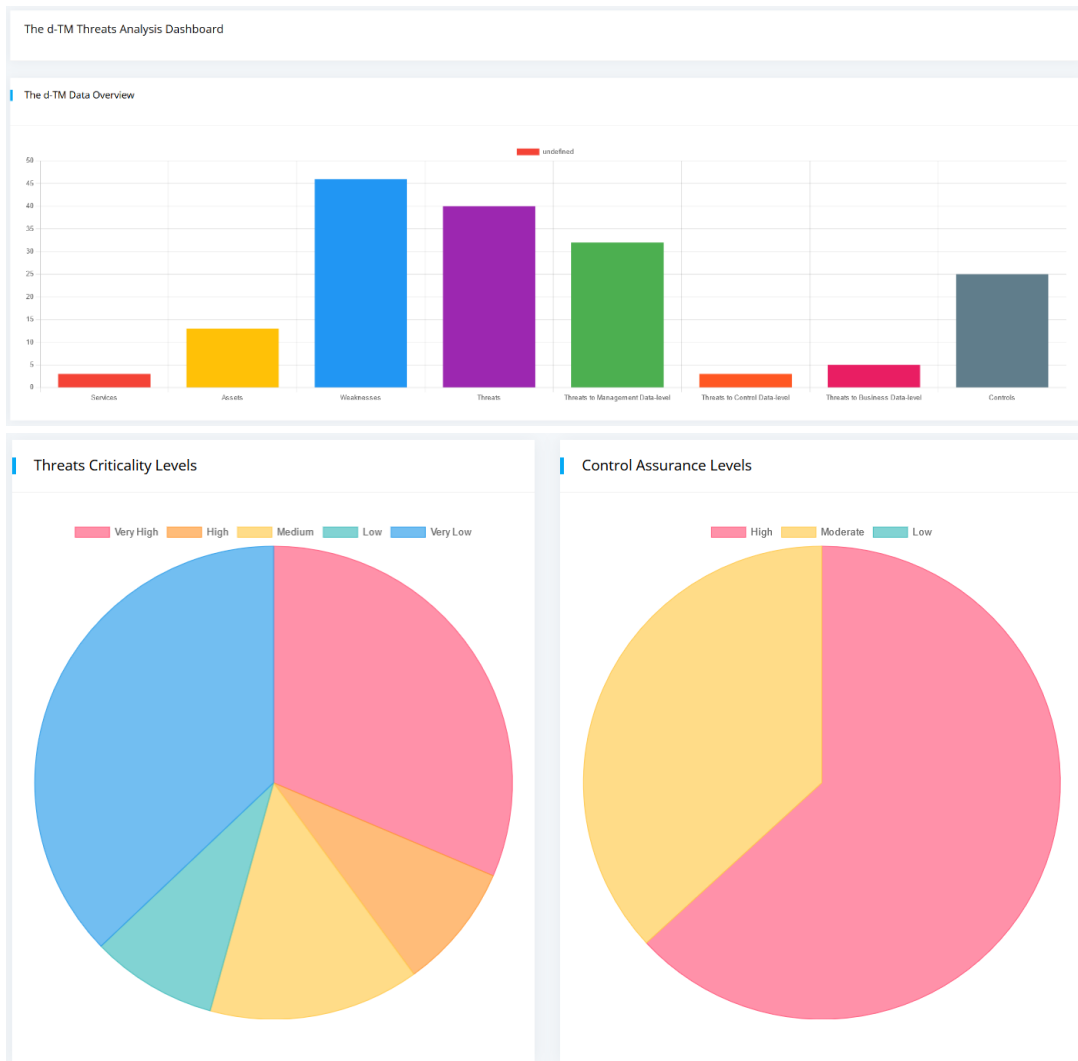


Figure 6.18. The d-TM dashboard of the case study threat analysis process.

The provided dashboard titled "The d-TM Threats Analysis Dashboard" offers a comprehensive visual representation of threat analysis metrics. The visual aids facilitate a quick grasp of the quantity and distribution of various parameters, supporting a more informed decision-making process.

- **Data Overview Bar Chart:**

- **Services:** The chart highlights three services. This sets the initial context, implying that the subsequent assets, weaknesses, threats, and controls are tied 'Patient e-service'.
- **Assets:** Representing thirteen assets which are the foundation upon which weaknesses and threats are identified.

- **Weaknesses:** The significant bar, corresponding to forty-six, indicates the weaknesses within the system. This higher count is a cause for concern as it determines potential points of exploitation.
- **Threats:** Out of the identified weaknesses, forty have been recognized as threats, suggesting that most weaknesses can be potentially exploited.
  - **Threats by Data Level:** A deeper dive reveals that most threats, thirty-two in total, target the Management-data level, making it the most vulnerable segment. The Control-data level and Business-data level have 3 and 5 threats, respectively. This distribution suggests that administrative and management practices of the system might be primary targets for malicious actors.
- **Controls:** The introduction of twenty-five controls indicates proactive measures taken to mitigate identified threats. Considering that there are forty threats and twenty-five controls, it is evident that some controls address multiple threats, underscoring their comprehensive nature.

- **Threats Criticality Pie Chart:**

The pie chart offers a color-coded representation of threat criticalities. Most threats fall within the ‘Very High’ to ‘High’ range, highlighting the urgency and significance of the potential risks. This distribution emphasizes the criticality of the identified weaknesses and the importance of deploying the necessary controls effectively.

- **Control Assurance Pie Chart:**

This chart provides an assurance breakdown for the identified controls. A substantial portion of the controls boasts a "High" assurance level, offering confidence in their efficacy. The remainder fall into the "Moderate" category, suggesting they might require further analysis to enhance their effectiveness.

In summary, this dashboard provides an intuitive overview of the threat landscape within the d-TM process. It underscores the need for robust countermeasures, especially at the Management-data level, given its heightened weakness. The provided controls, with their high assurance levels, signify a proactive approach towards threat mitigation, but the presence of a large number of threats with high criticality demands continuous monitoring and adaptation.

## Feature 2 Produce Visual Reports

This feature enables the d-TM platform to generate visual reports of each activity of the threat analysis. These reports are important to business stakeholders to share with relevant teams for informed decisions and collaboration. Figure 6.19 represents a sample of some generated reports for threat analysis. The visual reports can be produced for the following processes:

- Weakness identification process
- Threat identification process
- Criticality identification process

- Control identification process
- Control assurance identification process

Every report produced by the d-TM processes compresses a holistic understanding of the threat analysis, integrating elements such as asset identification, data levels, phases, critical threats, and assured controls.

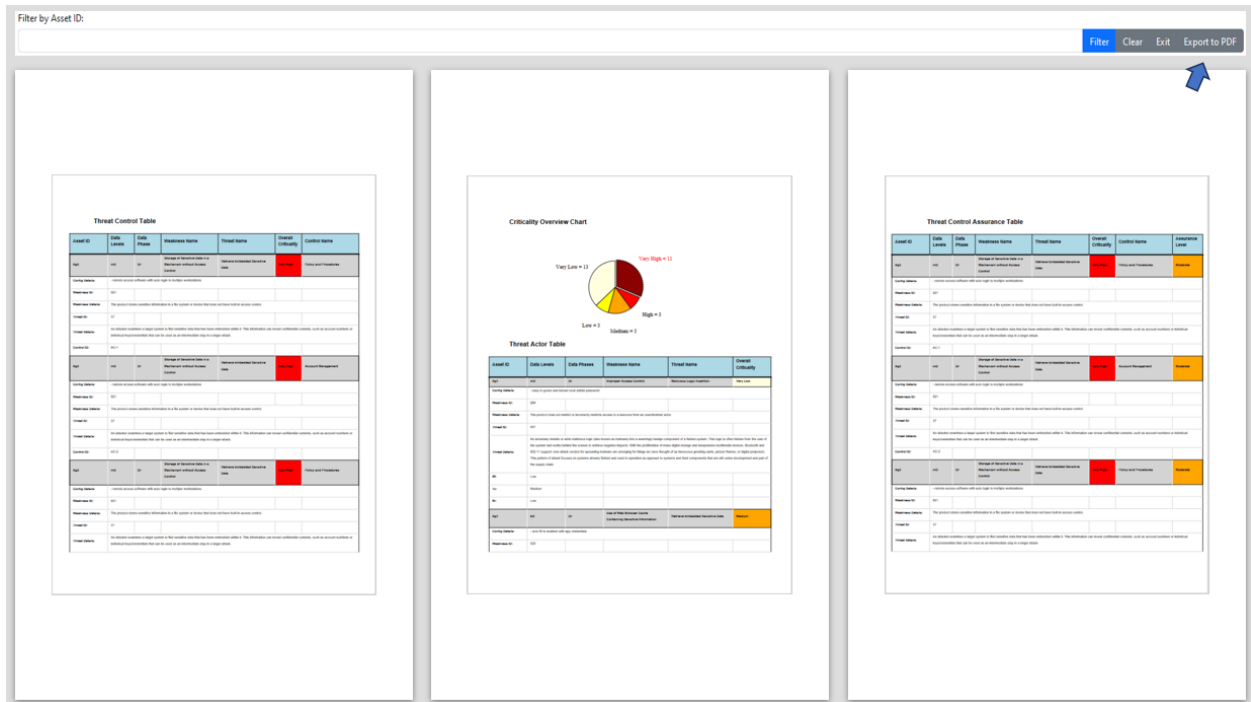


Figure 6.19. The d-TM visual reports samples for the case study threat analysis process.

## 6.4.3 Discussion

This section presents the observations gained after the implementation of the d-TM automation in the healthcare case study scenario. The use of d-TM within the examined scenario has a significant opportunity. The case study went over activities as follows:

- **Data and Business understanding**

The d-TM concept is based on data, and understanding running organizational data is crucial to the threat analysis process. The process consists of four activities, *beginning with understanding the business functions and its surrounding digital infrastructure*. The case study is a healthcare organization that utilizes digital services for day-to-day hospital operations. The process identifies that they rely mainly on the implemented health information system(HIS) application, which is called Careware. The careware is categorized by the d-TM as an extremely critical service to the scenario business continuity. In terms of underlying infrastructure supporting careware application, the process identifies thirteen assets classified to the d-TM five threat layers as follows: Agent(3), Network(5), Compute(2), Application(2), and Storage(1).

As d-TM is looking after data running the case study, ***the second process is to classify collected data from the previous process to the d-TM data level and phases***. This process assesses the collected information, config files and designs to find the proper mapping to the d-TM data concept. The process outcomes determined each data level and its phases for each asset; management(mD) and business data(bD) are the most common data levels identified, while data at rest(Dr) and data in transit(Dt) in terms of data phases.

Additionally, *Data Flow Diagrams (DFDs)* are used as visual representations to illustrate the connection between Careware applications and organization users. The diagram is employed to facilitate the understanding of how the underlying technology interacts with data, either directly or indirectly. The actor scenario outlined in the Data Flow Diagram (DFD) indicates that business users depend on hospital computers that have web browsers to act as agents for accessing the HIS application. In this scenario, the agent representing the business user initiates the request. The data then passes through multiple network components before getting to the careware application, which is hosted on two computing devices (an application server and a database server) and works with locally connected storage.

- **Weakness and Critical Threats**

As business logic is understood in the previous stage, ***identifying data levels, phases, weaknesses, and critical threats is the third process in the d-TM activities***. The case study is suffering from attacks that target its critical services, in particular, the HIS application. Attackers are considering infrastructure, system operators, and end users as a target to break through their defences. The threat analysis process is designed to evaluate information and diagrams collected to identify weaknesses and impose critical threats. The evaluation of the case study has produced the following information:

- **Weaknesses:** A total of forty-six weaknesses are identified across thirteen evaluated assets, mainly targeting management data. “Improper Access Control”, “Storage of Sensitive Data in a Mechanism without Access Control”, and “Cleartext Transmission of Sensitive Information” are considered the most identified weaknesses across critical service data assets.
- **Threats:** A total of forty threats are determined based on forty-six identified weaknesses. Most threats are targeting Ag<sub>1</sub>, Ag<sub>3</sub>, and Net<sub>4</sub>. The process identified several threats, but the most common threats included “Retrieve Embedded Sensitive Data”, “Interception, Application Fingerprinting”, and “Password Brute Forcing”.
- **Criticality to Data:** A total of eleven “very high” impact threats were identified that target Ag<sub>3</sub>, Net<sub>4</sub>, and App<sub>1</sub>. where threats are targeting mainly two data levels, which are management and business data levels. Furthermore, the data is at risk in two phases, which are at transit and while at rest. Moreover, three threats are classified as “high” impacts that are applicable to Cmp<sub>1</sub> and App<sub>2</sub>. However, the rest of the threats range from “medium” to “very low” impact.

In overall, this process identified significant weaknesses inside the organization, suggesting a possible unintentional oversight regarding the security of personal computers at the hospital, which are used by both normal users and system administrators. There seems to be a tendency to underestimate the critical need to implement security hardening measures for both networks and systems. From a technical standpoint, personal computers, which are often used for business-related and administrative functions, are now running software that has not been subjected to thorough examination. This software includes web browsers and terminal access tools. This exposes business data to potential risks, irrespective of its stage in the lifecycle, including creation, processing, or storage.

The system administrator's workstation presents a weakness that is of special significance. It was found that terminal access software is used with saved credentials for crucial network assets without proper access restrictions. If these workstations were to be hacked, it may potentially result in unauthorized access as a consequence of the credentials being exposed. In a comparable manner, network devices have been configured to use non-secure services for both access and data transmission, exposing them to flaws associated with Man-in-the-Middle (MITM) attacks. Furthermore, our assessment revealed that the Careware application and its hosting servers use a basic authentication mechanism. The threats that have been found possess the potential to significantly impact the integrity and security of the data discussed in the case study, regardless of whether it is being transmitted, processed, or stored.

- **Threat Mitigation**

The previous activities of the d-TM identify weaknesses and critical threats to careware applications. *Threat mitigation and assurance is the last activity in the d-TM process*, where suitable controls and assurance to threats are determined. This process mandated the implementation of determined controls in the case study to mitigate critical threats. Additionally, the controls are evaluated using the d-TM criteria to determine its assurance level to mitigate threats. The following illustrates the d-TM threat mitigation and assurance outcomes.

- **Controls:** A total of *twenty-five* controls are determined for *fourteen* critical threats. The controls that are identified in the process target “very high” and “high” classified threats; the controls are varied and majorly focus on account management, encryption, and organizational security policy enforcement.
- **Control Assurance:** The control examination highlights a total of *twelve* controls out of *twenty-five* are rated as having a “High” assurance to mitigate the threat, and *seven* controls are rated “Moderate”.

The implementation of d-TM in the case study reveals that the data is susceptible to threats, especially as a result of the identified critical threats. Mainly, these threats focus on compromising the system's authentication mechanism, the utilization of network unsecured services, and deficiencies in security guidelines enforcement. The d-TM process provides examined security measures to assure the cybersecurity assurance of the case study.

- **The Automated d-TM Scalability**

It's essential to highlight how the tool is engineered to adapt and scale with organizational changes, such as the integration of new subsystems, services, and data types. The d-TM tool's architecture is built with flexibility in its foundation, allowing for seamless updates and expansions to accommodate evolving organizational structures and technological advancements. The d-TM tool is designed in a modular approach such as data collection, analysis, and mitigation module aka application. It enables the addition of new components without disrupting existing functionalities, ensuring that the tool remains effective and relevant as organizations grow and diversify. For example, the tool can adopt new services, assets or data without interfering with the existing threat analysis process. This change is seamlessly integrated and provides dynamism to the existing organizational profile. Furthermore, the tool is designed to adopt different organization personnel using a dedicated module called Authentication. This module allows the organization to granularly assign different roles and privileges to each personnel in the organization's threat analysis activity i.e., business personnel, operator, and analyst. Similarly, The tool exhibits notable flexibility and scalability in facilitating threat analysis across the three abstracted data type levels, as well as within individual data-level types. This adaptability empowers organizations to concentrate their efforts on specific data types as needed. However, while this targeted approach can mitigate risks associated with the focused data-level, it may inadvertently expose the organization to threats pertinent to other data levels. For example, prioritizing the assessment of business data-level threats might safeguard against risks directly impacting business operations, but could leave vulnerabilities unaddressed in the management data-level, such as potential asset takeovers. It is crucial to recognize that each data type plays a vital role in a thorough threat analysis and contributes significantly to the organization's overall security posture. Ensuring a holistic approach that encompasses all data levels is essential for maintaining comprehensive protection against diverse threats. This inherent scalability ensures that the d-TM tool can continuously provide comprehensive threat modelling and analysis, even as organizations undergo significant changes and expansions.

- **The d-TM Expert Opinion**

Expert opinions are primarily sourced from the third case study, which is considered suitable because it aligns with the adoption of the tool. Cases one and two were conducted long before the tool was created. Moreover due, to time constraints and a lack of survey participants, from the company a survey couldn't be conducted. Hence only case study three was fully explored. The expert opinion is part of the d-TM evaluation process, these individuals possess extensive experience in networks, systems, and cybersecurity within diverse organizational contexts. these experts contributed to a comprehensive evaluation of the tool across multiple criteria, affirming its effectiveness, automation, and usability. Their feedback was instrumental in affirming the d-TM tool's strengths in identifying and mitigating threats effectively and efficiently, as well as its adoptability in integrating new organizational data and systems. The depth and breadth of their professional experience were essential in critically assessing the tool's capacity to streamline threat

modelling processes and enhance organizational security postures. The evaluation is conducted by five experts using a set of questions as attached in Appendix A. The outcome of this evaluation is summarized in Table 6.11.

<b>Criteria 1: Strengths and Weaknesses</b>	<b>Expert Responses</b>
1- How effective do you find the d-TM in identifying and mitigating threats across various organizational data types?	<ul style="list-style-type: none"> <li>• Very Effective: 80% (4 out of 5 respondents)</li> <li>• Somewhat Effective: 20% (1 out of 5 respondents)</li> </ul>
2- Can you quantify the coverage of organizational assets by the d-TM in your threat analysis? (e.g., percentage coverage)	<ul style="list-style-type: none"> <li>• More than 90%: 60% (3 out of 5 respondents)</li> <li>• 76% - 90%: 40% (2 out of 5 respondents)</li> </ul>
3- Have you encountered instances of false positives or negatives in threat detection with d-TM? Please provide estimates.	<ul style="list-style-type: none"> <li>• Rarely: 80% (4 out of 5 respondents)</li> <li>• Occasionally: 20% (1 out of 5 respondents)</li> </ul>
4- How many threats were identified from three levels of data, i.e., data --management, control ...? (open question)	<ul style="list-style-type: none"> <li>• Respondents reported identifying an average of 15-20 significant threats across data management, control, and other levels.</li> </ul>
5- How effective are the threat assessment layers, i.e., agent, network, ..etc?	<ul style="list-style-type: none"> <li>• Very Effective: 60% (3 out of 5 respondents)</li> <li>• Somewhat Effective: 40% (2 out of 5 respondents)</li> </ul>
6- How useful to measure the level of assurance for the d-TM proposed Control?	<ul style="list-style-type: none"> <li>• Very Useful: 80% (4 out of 5 respondents)</li> <li>• Useful: 20% (1 out of 5 respondents)</li> </ul>
<b>Criteria 2: Level of Automation and Efficiency</b>	
1- To what extent does the d-TM tool automate the threat management process in your experience? (Scale from 1 to 5)	<ul style="list-style-type: none"> <li>• 5 (Fully Automated): 80% (4 out of 5 respondents)</li> <li>• 4: 20% (1 out of 5 respondents)</li> </ul>
2- How significant was the reduction in human effort after implementing the d-TM tool in your threat analysis processes? (Quantify in terms of percentage reduction)	<ul style="list-style-type: none"> <li>• More than 90%: 40% (2 out of 5 respondents)</li> <li>• 76% - 90%: 60% (3 out of 5 respondents)</li> </ul>
<b>Criteria 3: Adoption of Open Intelligence</b>	
1- How do you assess the usefulness of adopted standards and open intelligence sources that are integrated into the d-TM tool in threat management?	<ul style="list-style-type: none"> <li>• Very Useful: 80% (4 out of 5 respondents)</li> <li>• Useful: 20% (1 out of 5 respondents)</li> </ul>
<b>Criteria 4: Adoptability of the d-TM tool</b>	
1- What has been the learning curve for new users adopting the d-TM tool in your organization? (Scale from Easy to Difficult)	<ul style="list-style-type: none"> <li>• Easy: 60% (3 out of 5 respondents)</li> <li>• Moderate: 40% (2 out of 5 respondents)</li> </ul>
2- Are the manual documentation and support provided with the d-TM tool sufficient for effective use and troubleshooting?	<ul style="list-style-type: none"> <li>• Very Sufficient: 60% (3 out of 5 respondents)</li> <li>• Sufficient: 40% (2 out of 5 respondents)</li> </ul>

3- How effective to visualise the data assets dependency using DFD of d-TM in understanding organizational infrastructure?	<ul style="list-style-type: none"> <li>• Very Effective: 80% (4 out of 5 respondents)</li> <li>• Somewhat Effective: 20% (1 out of 5 respondents)</li> </ul>
4- How difficult to generate the threat register using d-TM?	<ul style="list-style-type: none"> <li>• Easy: 80% (4 out of 5 respondents)</li> <li>• Moderate: 20% (1 out of 5 respondents)</li> </ul>
5- How do you benefit from using d-TM compared to previous practice? (open question)	<ul style="list-style-type: none"> <li>• All respondents highlighted significant improvements in threat detection, analysis efficiency, and overall cybersecurity posture as key benefits.</li> </ul>

Table 6.11. Overview of d-TM tool feedback.

The table reflects a predominantly positive reception of the d-TM tool among the expert respondents, with a particular emphasis on its effectiveness, automation, and integration of open intelligence.

**Criteria 1: Strengths and Weaknesses** The majority (80%) of experts rate the d-TM as very effective in threat identification and mitigation, suggesting a high level of confidence in the tool's capabilities. With 60% of respondents also reporting more than 90% coverage of organizational assets, the d-TM tool demonstrates a comprehensive reach within organizational infrastructures. A noteworthy 80% of experts rarely encountered false positives or negatives, indicating the tool's accuracy. However, a small percentage did occasionally encounter inaccuracies, highlighting a potential area for improvement. The detailed identification of threats across multiple data layers suggests that the d-TM tool offers a nuanced analysis, which is critical for thorough cybersecurity assessments.

**Criteria 2: Level of Automation and Efficiency** The level of automation is highly rated, with 80% of experts finding the d-TM tool to be fully automated. This underscores the tool's efficiency in streamlining threat management processes, reducing the need for manual intervention, and thus minimizing the potential for human error. The significant reduction in human effort, with 60% of respondents noting a 76% - 90% reduction, further underscores the d-TM tool's value in optimizing resource allocation.

**Criteria 3: Adoption of Open Intelligence** The high utility of open intelligence sources within the d-TM tool, as reported by 80% of experts, suggests that the tool is adept at incorporating and leveraging external security standards and intelligence in its threat management processes.

**Criteria 4: Adoptability of the d-TM tool** The adoptability of the d-TM tool is positively received, with the majority of users finding it easy to moderate to learn. This points to a user-friendly interface and suggests that the tool can be integrated into organizations without an extensive learning curve. The sufficiency of the manual documentation and support, as rated by respondents, is critical for ensuring that users can effectively utilize the tool and troubleshoot issues independently, reinforcing the tool's practical value.

Furthermore, figure 6.20 below presents a bar chart representation of the expert responses on various criteria based on the data provided. The green bars indicate a higher effectiveness,



automation, usefulness, or ease, as opposed to the blue bars which indicate a lower score in each respective category.

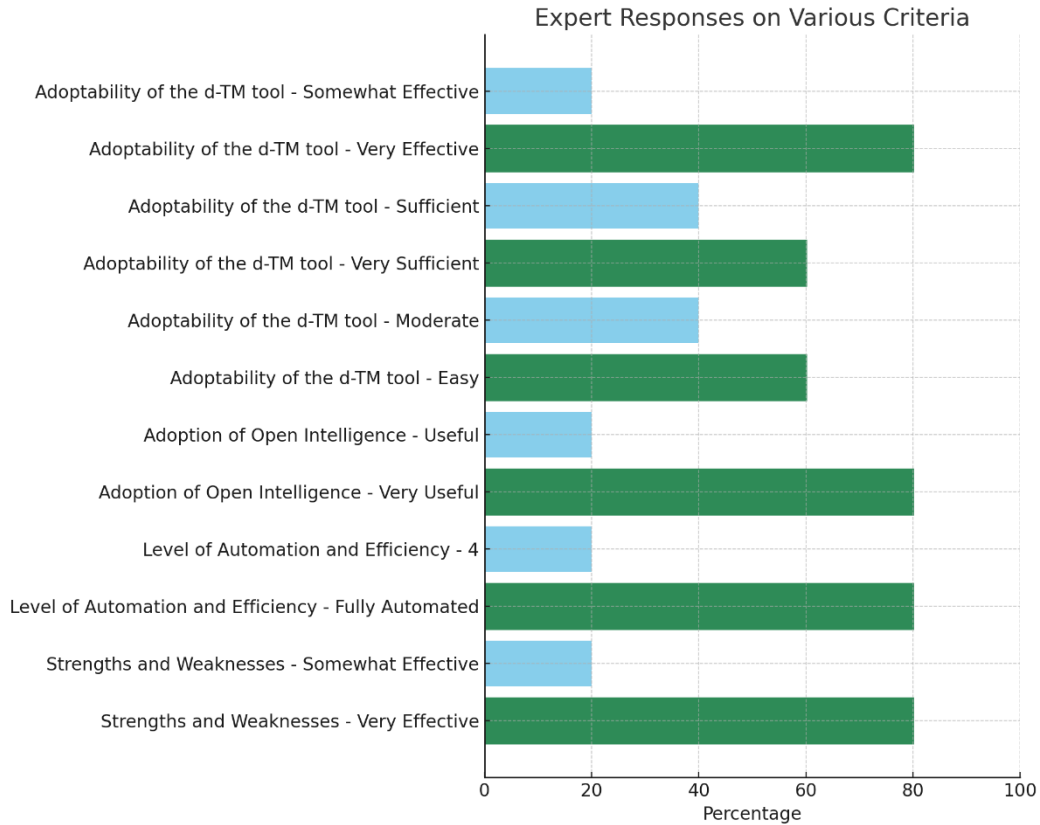


Figure 6.20. The expert opinion feedback.

In summary, the analysis of the table indicates that the d-TM tool is highly regarded in terms of its efficiency, effectiveness, and usability. The findings suggest that the tool's design and capabilities are well-aligned with the needs of cybersecurity professionals, offering a scalable and adaptable solution for comprehensive threat management. However, the occasional incidence of false detections and the moderate difficulty reported by some users provide areas for further refinement to enhance the tool's precision and user experience.

## 6.5 Comparison with the other works

This section provides an overview of the evaluation summary of the d-TM, focusing on its similarities to previous works and the additional value it brings. The objective is to identify the strengths and weaknesses of the concept and its application. The comparison evaluated the assessed case situations with relevant literature.

- **Application of d-TM**

Threats to an organization's infrastructure may be comparable, particularly if they use the same digital paradigm, such as healthcare or cloud-enabled technologies. The research discussed three case scenarios for different industries that aim to evaluate d-TM concept applicability over a wider

range of industries. The cloud case study assessed by the research reveals two weaknesses that are highlighted by The OWASP "Top 10 cloud risks"(Cisco, 2015), which are "User Identity Federation" and "Infrastructure Security". Furthermore, the Denial-of-Service(DoS) attack is discovered in the same scenario and has been recognized by other scholars' research, such as (Alexander and Wang, 2019), (Alouffi *et al.*, 2021),(Giannoutakis *et al.*, 2020). Additionally, the three case studies discussed the threats to organization identity management, which are also revealed by the assessment, which is similarly determined by (Lakhanpal, 2020) (Ganiga *et al.*, 2020) in his research, where he stated that account takeovers are a frustrating security issue in distributed computing since they are usually difficult to detect. In the healthcare scenario, (Newaz *et al.*, 2022) highlights "interception" and "unauthorized remote access" as top attacks that could be exploited due to software vulnerability or malicious software. Revealed attacks and caused weakness are also identified in the d-TM evaluation case study. Furthermore, the consequences of insufficient account management and intermediate network device misconfiguration are even more difficult to trace and mitigate. However, most of the similarities in detected threats are related to applications or systems, not overall digital assets interacting with data. Nevertheless, the d-TM analysis was also able to uncover additional threats that were not covered by others, which are related to data and underlying infrastructure, such as Agent(Ag) used to access applications or administrative tasks and Network(Net) misconfiguration in our scenarios. d-TM also provides superior value due to provided features, including the layered attack approach, enhanced DFD presentation, and mapping threat analysis outcomes to industry standard catalogues, weaknesses based on MITRE CWE, threats described by MITRE CAPEC, and NIST 800-53 controls for mitigation.

- **Automation of d-TM**

The automation capability is one of the key strengths of the d-TM model. Unlike many traditional threat modelling approaches that heavily rely on manual processes, d-TM modernizes and advances threat analysis using a dedicated automation platform. This platform is designed to automate the d-TM threat analysis process. The foundation of the d-TM process relies on three features, data collection, threat analysis, and automation features.

**Data Collection:** The d-TM platform provides crucial features to existing models by actively engaging technical and business stakeholders through the data collection process, ensuring that essential information is gathered efficiently. Where it guides platform users through a sequence of input forms designed to extract critical insights about the organization's business processes, and digital infrastructure which are vital for a comprehensive threat analysis.

**Threat Analysis:** Once the information is gathered, the tool guides the security analyst over a sequence of linked forms that are mapped to the d-TM threat analysis process and provides required directions and insights to identify weaknesses, threats, and suitable controls. The platform is empowered by an automated algorithm to identify threats criticality and control efficiency towards determined controls.

**Automation:** There are multiple tools that are used by industry and scholars to aid in threat modelling process, these tools could open-source or commercial. Table 6.12 represents a

comparison of threat modelling automation aspects between d-TM platform and some of common threat modelling tools in the industry. However, In the comparative table, the tools under study are acknowledged as some of the leading threat modelling tools currently embraced by the industry, as noted by (Kirvan, 2023). *Please note*, the evaluation of these tools was thoroughly carried out by examining their official documentation and empirical evaluations carried out within the researcher's lab environment. The main goal of this investigation is to find similarities and recognize constraints to the d TM tool for future advancement, while still acknowledging the importance of other tools, in the field.

<b>Aspects\Tools</b>	<b>OWASP Threat Dragon</b> <i>(OWASP Threat Dragon   OWASP Foundation, 2023)</i>	<b>Microsoft Threat Modelling</b> <i>(Microsoft Threat Modeling Tool overview - Azure, 2022)</i>	<b>IriusRisk Threat Modelling</b> <i>(Threat Modeling Platform, 2023)</i>	<b>d-TM Data-Driven Threat Modelling</b>
<b>Use Case Understanding</b>	Technical Driven	Technical Driven	Technical Driven	Business and Technical Driven
<b>Focus</b>	Systems and Applications	Systems and Applications	Systems, Applications, and Business Data	Data-Driven
<b>Asset Details</b>	Basic Details (such as name, role, etc.)	Basic Details (such as name, role, etc.)	Basic Details (such as name, role, etc.)	Comprehensive Details (e.g., OS, IP, Dependencies, etc.)
<b>Threats Models used</b>	STRIDE	STRIDE	STRIDE	d-TM
<b>Threat Catalogue</b>	No (Manual entry)	No (Manual entry)	Yes – multiple catalogues, i.e., CAPEC	Yes - CAPEC
<b>Threat Root Cause</b>	No	No	Based on CWE	Based on CWE
<b>DFD Construction and Concept</b>	Manual Generated	Manual Generated	Manual Generated	Automatic based on Data Asset Dependency
<b>Threat Layers</b>	No (based on Expert Judgment)	No (based on Expert Judgment)	Asset based - Unguided (Based on Assets and Expert Judgment)	Defined and Guided (Five layers)
<b>Criticality</b>	Manual, Expert Judgment	Manual, Expert Judgment	Based on CIA and Ease of Exploitation	By d-TM Three Factors
<b>Mitigation</b>	Yes (Manual based on Expert Judgment)	No	Yes - Multiple Standards, i.e., NIST	Yes - NIST
<b>Control Assurance</b>	No	No	No	Yes, Driven by Three Factors

Table 6.12. Overview of d-TM tool to some existing tools.

- **Visualization in d-TM**

The visual output produced by d-TM, is a distinguishing feature that sets it apart from other threat modelling approaches. The d-TM platform provides data-centric insights for each process outcome. Each activity in the process of threat assessment is visualized in a comprehensive color-coded table that combines the current process outcome and any previously relevant identified data.

This thoughtful design allows users to generate tailored outputs for distinct aspects of the process, whether it is related to critical business services, infrastructure inventory, or identifying weaknesses and threats. Moreover, the d-TM platform is considered a step further with its automated produced Data Flow Diagrams (DFD). The diagrams provide an advanced way of visualizing how data moves within an organization. By representing data dependencies among various assets, the platform helps platform users to comprehensively understand the data asset interactions. Also, the d-TM platform also empowers users to create visual reports for each process. These reports serve as a powerful tool for sharing threat assessments with relevant stakeholders.

- **Usability of d-TM**

Usability is a significant aspect of the d-TM threat modelling approach; it places a strong emphasis on making the threat analysis leveraged by a user-friendly interface and easy-to-access implementation. The d-TM interface is designed to guide users through its process effectively. It ensures that technical stakeholders can provide the necessary information without requiring extensive knowledge of threat modelling. Furthermore, leveraged automation capabilities in d-TM reduce the burden of manual data collection and analysis, making threat modelling more time-efficient and less labour-intensive. This enhances usability for organizations with varying levels of technical expertise. Also, d-TM's usability extends to non-technical stakeholders, such as business decision-makers. The visual output, including DFDs, intensive insights tables, and produced reports simplifies the communication of threat analysis outcomes, fostering a shared understanding of cybersecurity risks.

# CHAPTER *SEVEN*: THE CONCLUSION

## 7.1 Introduction

The modern industry landscape is distinguished by a high reliance on data and digital technologies, with organizations increasingly viewing data as a strategic asset. This digital evolution has contributed to an increase in cyber risks, requiring proactive and efficient cybersecurity solutions. The Data-driven Threat Modelling (d-TM) approach is proposed in this thesis, which represents significant advancements in the field of threat modelling and cybersecurity assurance. By acknowledging the fundamental importance of cybersecurity in modern digital surroundings, where businesses are continuously at risk from cyberattacks that could disrupt their business continuity. A comprehensive review of existing threat analysis models and practices is conducted in this thesis, which identifies certain gaps and limitations, particularly in addressing data-related threats and evaluating the effectiveness of security controls. In this concluding chapter, we summarize our research findings and reveal the d-TM model's novel contribution that demonstrates its critical role in enhancing the field of threat modelling in particular focusing on data-driven threat analysis and cybersecurity assurance.

## 7.2 Fulfilling Research Questions

This section presents how the research questions, have been comprehensively addressed and fulfilled through the development and validation of the d-TM model. Each question is revisited with a concise summary of the methodologies employed and the key findings that contribute to the resolution of the stated inquiries.

### 7.2.1 RQ1: What are the existing gaps in the state of the art with regard to threat modelling approaches?

The exploration of existing threat modelling practices revealed several gaps, such as analysing and managing data-related threats. The d-TM model, through its innovative use of open intelligence sources, threat layers, actors, and automated analysis, directly addresses these gaps. With that, the d-TM model enhances the adaptability and efficiency of threat analysis and management practices, thereby offering a substantial advancement over traditional methods.

### 7.2.2 RQ2: What are the various types of organizational data that need to be analysed?

In response to the second research question, the d-TM model employs a comprehensive data mapping strategy that identifies and categorizes various types of organizational data. This strategy includes an in-depth analysis of business data, and operational data at any stage of its lifecycle, thereby ensuring a holistic approach to threat identification and analysis. This process not only meets the objective of enhancing organizational data understanding but also lays a foundational framework for effective threat modelling.

### 7.2.3 RQ3: What are the various types of organizational data that need to be analysed?

By employing a data-centric approach, the study distinguishes data into distinct levels and phases, ensuring a thorough examination of weaknesses and threats across all types of data assets. This methodology not only highlights the critical data types within an organization but also enhances

the precision of threat analysis. In detail, the approach involves a detailed process of data mapping and categorization, which is foundational to the model's effectiveness. This process begins with an extensive assessment of the organizational data ecosystem, identifying various data types in three abstraction levels, such as business data, control data, and management data. Furthermore, the model analyses how these various data types interact within the organizational ecosystem, including their flow through different systems and networks. This holistic analysis is facilitated by the model's use of data-driven Data Flow Diagrams (DFDs), which visually represent the movement and interaction of data across the organization. In summary, this abstract categorization ensures that all potential data sources are considered, providing a holistic view of the organization's data landscape, and thereby fulfilling the research question's objective.

#### **7.2.4 RQ4: How can the identified threats be prioritized to determine the appropriate level of assurance for effective mitigation?**

The d-TM model introduces a novel methodology for the prioritization of threats, incorporating Three factors, i.e., Business-as-a-target (Bt), Threat complexity (Tc), and Business Impact (Bi). These factors enable d-TM to integrate business and technological understanding into the prioritization process. This methodology enables organizations to allocate resources effectively, focusing on mitigating the most critical threats first. This approach not only optimizes resource utilization but also ensures that the level of security assurance is aligned with the organization's risk tolerance and business objectives.

#### **7.2.5 RQ5: To what extent can threat modelling be automated to enhance its effectiveness and facilitate wider adoption for overall security assurance?**

The d-TM model presents a significant advancement through the development of an automated tool. This tool streamlines the threat analysis and management process, including the whole Four activities, i.e., data collection, data analysis, threat analysis, and threat mitigation, making it more efficient and adaptable to a wider range of organizations. The automation of threat analysis and the generation of actionable insights facilitate a significant feature of cybersecurity practices, enhancing the model's effectiveness and its adoption in real-world scenarios.

By systematically addressing each research question, this thesis contributes valuable insights and methodologies to the field of threat modelling and analysis. The development and validation of the d-TM model, supported by Three real-world case studies, underscore the model's effectiveness in enhancing organizational security assurance through innovative, data-driven threat analysis and mitigation strategies.

### **7.3 Fulfilling Research Objectives**

The proposed d-TM model has been developed and validated using three real-world case scenarios to meet the primary objective of this research to

*“Develop an innovative and automated data-driven threat modelling approach that empowers organizations to effectively address data-related cybersecurity threats for overall security assurance”.*

To ensure the research aim is satisfied, the objectives were specified as:

### **7.3.1 Enhance the understanding of organizational data**

The first objective is to empower organizations with a comprehensive understanding of business running data. this objective is fulfilled by the d-TM model as follows:

- **Comprehensive Data Mapping:** the model identifies and categorizes data types, locations, and dependencies across three levels of abstraction.
- **Asset Identification:** within the d-TM process, data assets are identified and classified into five asset types (Agent, Network, Compute, Application, and Storage). These assets represent the data lifecycle from user endpoint to data storage.
- **Visual Representation:** d-TM extends its analysis with visual representations of data assets and their relationship in perspective to data. The presentation is produced as a Data Flow Diagram (DFD). This visual representation leverages a comprehensive understanding of data flows and their interactions within the organization.

### **7.3.2 Develop an innovative data-driven threat analysis modelling approach**

The core objective was to create a novel data-driven threat analysis modelling approach that capitalizes on data to evaluate and strengthen cybersecurity assurance. This entailed the conceptualization and producing a model in which data stands paramount within the threat analysis procedure.

The d-TM presents a significant value to existing threat models due to the many features that empower the model. The model considers business processes and services as the initial reference point to analyse threats, which is overlooked by existing contributions. It also defined attack layers and data levels for the threat analysis. D-TM adds another excellent value compared to other models that keep attack surfaces to expert judgments, which could result in inconsistency in reproducing results or overlooking a necessary attack surface. As a result of business processes, services and layered enabled infrastructure understanding, d-TM translates this information to a data-driven DFD diagram that presents data types and phases of any infrastructure asset. A data-driven leveraged DFD provides an important advantage to any organization, traditional application, or as-set base DFD used in most research is lacking the feature of data level and phase. The threat analysis process considers weaknesses identification based on CWE KB, which is later used to determine implied threats to data; the threats also take advantage of CAPEC KB. d-TM provides a new way to identify the criticality of each threat, the criticality is determined using three necessary factors: Bt, Tc and Bi. These factors enable d-TM to integrate business understanding into the prioritization process. Lastly, identified critical threats are evaluated to determine suitable controls and their assurance of business objectives. None of the discussed research considers Three-data levels, data-oriented DFD, Three-Actors use cases, five attack layers (especially Agent), and common KB to analyse data threats. However, d-TM is based on data, so the identified risks are both generic and unique in our case. Furthermore, d-TM is leveraged with automation tools to enhance the applicability of the d-TM process for industry use. Table 7.1 shows a qualitative comparison of d-TM with existing threat models.



<b>Factors/Models</b>	<b>PASTA</b>	<b>STRIDE</b>	<b>d-TM</b>
<b>Threat Modelling Methodology</b>	Attack-Centric	Threat-centric	Data-Centric
<b>Threat Modelling Stages</b>	7	N/A	4
<b>Consideration of Business</b>	Yes	No	Yes
<b>Identification of Threats Criticality and Control</b>	Yes	No	Yes
<b>Data-enabled DFD</b>	No, System or Application	No, System or Application	Yes, Data, System and Application
<b>Process Automation</b>	No	Yes	Yes
<b>Identification of Attack Surface</b>	Attacks to: - Network - Compute - Application	Threats to : - Compute - Application	Threat to Data in : - User-agent - Network - Compute - Application - Storage
<b>Data Area of Focus</b>	Single category, three phases	Single category	Three categories: Management/Control/Business, Three phases.
<b>Threat Modelling Components</b>	- Business - Asset/Application - Motivation/ Scenario - Vulnerability - Control	- Asset - Vulnerability - Threat - Control	- Business/ Asset - Data/ Vulnerability - Threat - Control - Assurance

Table 7.1 Existing threat modelling to d-TM overview(Alwaheidi, Islam and Papastergiou, 2022).

### 7.3.3 Validate the effectiveness of the proposed threat modelling approach

To demonstrate the effectiveness and practicality of the proposed data-driven threat modelling (d TM) methodology we conducted three real-world case studies. These studies covered sectors, including an IT solutions provider, a supply chain integrated with cloud technology and a healthcare facility. Although each organization had functions, they all shared a common goal; to enhance their business revenue and ensure a flawless operational experience by implementing robust security protocols.

Chapter 6 explains how the d TM methodology was implemented in these world industrial contexts. It does not discuss the insights gained through the application of the d TM model to identify potential threats but also records the specific steps involved in the threat analysis process. The evaluation focused on two objectives. Firstly, it aimed to assess how suitable and relevant the d TM approach is, in settings. Secondly, it aimed to identify any challenges or pitfalls that may arise during hands-on deployment of the d methodology.

As depicted in Chapter 6, Figure 6.1 titled "d TM Deployed Case Scenarios" our selected case scenarios cover industries such as supply chain management, solutions service provisioning and

healthcare. In the two cases, the d-TM process was applied manually while the third case utilized the automated capabilities of the d-TM platform. These scenarios had their infrastructure setups ranging from cloud-based to on-premises systems. As a result, a detailed threat analysis was generated by the activity, which evaluated data assets identified weaknesses and highlighted emerging threats. It also suggested security controls to counter these threats.

Interestingly despite each organization's distinctiveness a comparative analysis of the results from these case studies revealed patterns. This emphasizes that there are threats and vulnerabilities faced by these entities demonstrating how relevant and effective the d TM approach is, in real-world scenarios. This empirical evidence supports the objective of this thesis; developing a data-driven threat analysis paradigm that enhances cybersecurity assurance.

#### **7.3.4 Develop a tool to automate the d-TM threat analysis process**

Threat modelling plays a role in helping organizations protect their environments. While there are frameworks for threat modelling many of them require significant manual effort making their implementation difficult. Introducing an innovative approach, the data-driven threat modelling (d TM) methodology emphasizes efficiency by using data to identify and counter cyber threats.

The research in chapter five highlights the focus on introducing a d-TM tool that is designed to automate the traditionally time-consuming and human-intensive process of threat modelling. This tool provides a comprehensive workflow that guides users through defining business objectives, services, and infrastructure. It also conducts threat evaluations. Helps develop strategic plans for mitigating risks. In addition to process guidance, the d TM tool generates reports that summarize its findings enabling organizations to make informed cybersecurity decisions based on data. The key features of the d-TM tool include:

- **Asset Identification:** Identifying services and valuable data assets.
- **Threat Analysis:** Detecting weaknesses and potential threats.
- **Mitigation Planning:** Suggesting controls to neutralize identified threats and evaluating their effectiveness.

The automation capabilities of the d-TM significantly reduce the time-consuming nature of threat analysis enabling focus on detecting existing risks. This automation goes beyond improving implementation process efficiency; it greatly enhances the cybersecurity stance of an organization. The foundational processes of the d TM Platform.

- **Data Collection:** The d-TM tool emphasizes on engaging stakeholders, during the data-gathering phase. Through user input forms it gathers insights about business operations and digital infrastructure laying the foundation for detailed threat analysis.
- **Threat Analysis:** Once data is collected the tool guides users through a process using interconnected forms that align with the d TM threat analysis methodology. Empowered with an innovative algorithm that assesses threat severity and ensures efficiency of control selection.

Automation tools can be comparable, the research in Chapter 6 'Table 6.11' provides a comparison between d-TM and well-known tools such as OWASP Threat Dragon, Microsoft Threat Modelling

and IriusRisk. This comparison is based on practical evaluation in the researcher's lab environment, this aims to find commonalities and limitations in the d-TM tool, without underestimating other tool's value. The assessment highlights unique features, in areas, including a focus on data, the use of automated Data Flow Diagrams (DFD), and ensuring control. Additionally, the research discusses the applicability, usability, and visualization aspects. As a result, this empirical analysis supports the objective of the thesis, which is centred around creating an instrument for analysing threats based on data. This tool does not enhance the usefulness of the d model but also provides a valuable addition, to the existing variety of threat modelling tools.

In conclusion, the primary research objective of formulating an innovative and automated data-driven threat modelling methodology, designed to empower organizations in efficiently addressing data-centric cybersecurity threats, has been accomplished. This fulfilment has been realized through the successful achievement of the four underlying goals. Firstly, by extending the comprehension of organizational data; secondly, through the creation of an innovative data-driven threat analysis modelling approach; thirdly, by validating the efficacy of the proposed threat modelling methodology; and lastly, by engineering a tool that streamlines and automates the d-TM threat analysis process. These achievements collectively underscore the potential of this research in strengthening overall security assurance for organizations, marking a significant contribution in the realm of cybersecurity.

#### **7.4 The d-TM observed limitation**

The assessment of the d-TM is carried out by using three real-case scenarios. In the context of implementing the d-TM automation platform, several observations have been made about its limits. It is important to emphasize these limitations to overcome them in future research efforts. The limitations outlined below are especially relevant to the d-TM platform, which serves as the practical result of this research effort.

- The implementation of the d-TM, specifically with the identification of assets, maybe a time-consuming process, particularly for organizations of medium to large size. The process delay has the potential to impede the overall efficiency of the platform's application. To streamline this particular component and optimize customer satisfaction, potential future versions of the platform may consider including an automated discovery functionality. This proposed feature would include the implementation of an automated scanning system that can efficiently gather the necessary information pertaining to the organization's infrastructure. The optimisation of usability and applicability of the platform might be achieved by minimising the human work required for asset identification, which would be especially beneficial for organizations with complex technological ecosystems.
- The present architecture of the d-TM platform is tailored to address the specific requirements of different organizations in the context of threat analysis. Although this constraint may not impose significant limitations on organizations using the platform for internal reasons, it may provide a restriction for security solution providers aiming to achieve a broader reach. To tackle this issue, one possible approach for enhancement might include the integration of multi-scheme database functionality inside the platform. This proposed improvement will expand the platform's ability to offer services to a wider range

of clients, effectively meeting the needs of security solution providers that operate in various organisational settings.

- One notable observation pertaining to the platform is the complexities involved in the process of picking suitable weaknesses, threats, and control identifications. To successfully do this procedure, it is important that users have a thorough understanding of the MITRE CWE and CAPEC catalogues, in addition to the NIST controls. Due to the constantly changing and extensive nature of these catalogues, as well as the extensive number of items requiring evaluation, the technicality of this selection process has the potential to impact progress negatively. To overcome this difficulty, it may be beneficial to investigate the incorporation of a machine-learning system. The use of this algorithmic methodology has the capacity to optimize the selection procedure by displaying a determined subset of relevant elements, thereby reducing the complicated nature of the available options. This has the potential to enhance the decision-making process for security practitioners who are using the platform.

The d-TM platform offers a potentially effective method for conducting threat analysis. However, the identified shortcomings shed light on specific aspects that need further improvement and development. By acknowledging and mitigating these constraints, the platform can broaden its scope, improve its usability, and optimize its efficacy in supporting cybersecurity initiatives.

## 7.5 The d-TM Future work

This thesis represents a significant step forward in the threat modelling domain, it is essential to acknowledge that the field of cybersecurity is ever-evolving. Future research actions could explore the integration of machine learning and artificial intelligence to enhance threat prediction and mitigation further. Moreover, the research can explore the scalability of d-TM, especially in medium to large-sized organizations. Usability is another significant factor for any tool for wide adoption, that side; also needs attention in future work. The development of the d-TM features is continuous and will be carried on achieving the following advancements:

1. **Automated Asset Discovery:** The initial implementation of d-TM, particularly in asset identification, has been observed to be time-intensive, especially for medium to large-scale organizations. This time commitment could potentially hinder the overall efficiency and user experience. To improve this, it is proposed that future enhancement incorporate an automated discovery functionality. This would entail developing a feature that can autonomously scan and gather pertinent details about an organization's infrastructure. By reducing the manual effort needed for asset identification, the platform's usability and applicability would be significantly enhanced, especially beneficial for organizations with complicated technological landscapes.
2. **Integration of Multi-Scheme Database Functionality:** The existing architecture of the d-TM platform is designed to conduct threat analysis for a single organization. While this provides a limitation, it may act as a bottleneck for security solution providers or multi-organizational entities aspiring for wider adaptability. A prospective advancement is to embed multi-scheme database capabilities within the platform, broadening its service scale

and ensuring it caters to diverse organizational matrices, thus making it more flexible for security solution providers.

3. **Machine-Learning-Driven Selection Process:** A significant observation reveals complexities in selecting suitable weaknesses, threats, and controls. This process demands users to have an in-depth knowledge of MITRE CWE, CAPEC catalogues, and NIST controls. Given the evolving and vast nature of these repositories, the technical nature of this selection can potentially impede efficiency. A promising opportunity is to delve into integrating machine learning algorithms. Such an integration would streamline the selection procedure, presenting users with a focused list of choices. This refinement stands to elevate the decision-making experience for security practitioners, making the platform more intuitive.
4. **Integration of Human-Factor:** Threat modelling and management have traditionally been technology, threats, or data-centric driven, often overlooking the focus on human factors in the entire threat analysis process. However, in today's complex digital ecosystem, where human-technology interactions are increasingly connected, there is a necessary need to reconsider this narrow focus. The approach to integrating the human-factor into the d-TM will be initiated by identifying key human factors such as motivation, knowledge, context, and privilege. By doing so, it will offer a granular understanding of how these variables influence security postures within organizations. The integrated elements will recognize the multifaceted nature of human behaviour in cybersecurity risk—where motivation encompasses the drivers behind actions, knowledge pertains to the understanding of both system and security protocols, context relates to the specific organizational roles and environments, and privilege defines the levels of system access granted.

In summation, while the d-TM platform has carved a niche for itself in threat analysis, the highlighted limitations present opportunities for enhancement. Addressing these will not only refine the platform but will also position it as a more comprehensive tool in the realm of cybersecurity. The future beckons with promising avenues to make the d-TM platform more robust and universally adaptable.

## 7.6 Summary

In the concluding chapter of this research, the focus has been on revisiting and emphasizing the achievements of the study in line with its initial objectives. The study began with the identification of an existing need to address data-related cybersecurity threats in a more efficient and effective manner. This led to the formulation of the primary objective that aims to develop an innovative and automated data-driven threat modelling approach.

The approach presented in this thesis, known as Data-driven Threat Modelling (d TM) provides a thorough strategy, for dealing with data-related threats and enhancing cybersecurity assurance. By emphasizing the importance of data, weakness, threats, control evaluation and automation. D-TM leverages organizations with the tools and insights to navigate the ever-changing field of cybersecurity. As threats continue to evolve it is crucial that our approaches to threat modelling evolve as well. The development of d TM demonstrates our commitment to achieving excellence, in cybersecurity.

# Appendix A: Set of d-TM Evaluation Questions

## Criteria 1: Strengths and Weaknesses

1- How effective do you find the d-TM in identifying and mitigating threats across various organizational data types?

- Very Ineffective
- Somewhat Ineffective
- Neutral
- Somewhat Effective
- Very Effective

2- Can you quantify the coverage of organizational assets by the d-TM in your threat analysis? (e.g., percentage coverage)

- Less than 25%
- 26% - 50%
- 51% - 75%
- 76% - 90%
- More than 90%

3- Have you encountered instances of false positives or negatives in threat detection with d-TM? Please provide estimates.

- Frequently
- Occasionally
- Rarely
- Once
- Never

4- How many threats were identified from three levels of data, i.e., data --management, control ...? (open question)

- .....

5- How effective are the threat assessment layers, i.e., agent, network, ..etc?

- Very Ineffective
- Somewhat Ineffective
- Neutral
- Somewhat Effective
- Very Effective

6- How useful to measure the level of assurance for the d-TM proposed Control?

- Not Useful
- Somewhat Useful
- Neutral

- Useful
- Very Useful

### **Criteria 2: Level of Automation and Efficiency**

1- To what extent does the d-TM tool automate the threat management process in your experience? (Scale from 1 to 5)

- 1 (Minimal)
- 2
- 3
- 4
- 5 (Fully Automated)

2- How significant was the reduction in human effort after implementing the d-TM tool in your threat analysis processes? (Quantify in terms of percentage reduction)

- Less than 25%
- 26% - 50%
- 51% - 75%
- 76% - 90%
- More than 90%

### **Criteria 3: Adoption of Open Intelligence**

1- How do you assess the usefulness of adopted standards and open intelligence sources that are integrated into the d-TM tool in threat management?

- Not Useful
- Somewhat Useful
- Neutral
- Useful
- Very Useful

### **Criteria 4: Adoptability of the d-TM tool**

1- What has been the learning curve for new users adopting the d-TM tool in your organization? (Scale from Easy to Difficult)

- Very Easy
- Easy
- Moderate
- Difficult
- Very Difficult

2- Are the manual documentation and support provided with the d-TM tool sufficient for effective use and troubleshooting?

- Insufficient

- Somewhat Sufficient
- Neutral
- Sufficient
- Very Sufficient

3- How effective to visualise the data assets dependency using DFD of d-TM in understanding organizational infrastructure?

- Very Ineffective
- Somewhat Ineffective
- Neutral
- Somewhat Effective
- Very Effective

4- How difficult to generate the threat register using d-TM?

- Very Easy
- Easy
- Moderate
- Difficult
- Very Difficult

5- How do you benefit from using d-TM compared to previous practice? (open question)

- .....



## References

- Abdulghani, H.A. *et al.* (2019) 'A Study on Security and Privacy Guidelines, Countermeasures, Threats: IoT Data at Rest Perspective', *Symmetry*, 11(6), p. 774. Available at: <https://doi.org/10.3390/sym11060774>.
- Aftab, S. *et al.* (2018) 'Empirical Evaluation of Modified Agile Models', *International Journal of Advanced Computer Science and Applications (ijacsa)*, 9(6). Available at: <https://doi.org/10.14569/IJACSA.2018.090641>.
- Aghdash, S.A. *et al.* (2021) 'Concepts and Applications of Action Research in Improving the Performance of the Health System: A Guide for Managers', *Depiction of Health*, 12(3), pp. 273–285. Available at: <https://doi.org/10.34172/doh.2021.27>.
- Alexander, C.A. and Wang, L. (2019) 'Cybersecurity, Information Assurance, and Big Data Based on Blockchain', in *2019 SoutheastCon. SoutheastCon 2019*, Huntsville, AL, USA: IEEE, pp. 1–7. Available at: <https://doi.org/10.1109/SoutheastCon42311.2019.9020582>.
- Alexander, F. (2017) '10 types of data that change your business game - IBM Business Analytics Blog'. Available at: <https://www.ibm.com/blogs/business-analytics/10-types-of-data-change-business/> (Accessed: 22 February 2020).
- Ali Alatwi, H. and Morisset, C. (2022) 'Threat Modeling for Machine Learning-Based Network Intrusion Detection Systems', in *2022 IEEE International Conference on Big Data (Big Data). 2022 IEEE International Conference on Big Data (Big Data)*, pp. 4226–4235. Available at: <https://doi.org/10.1109/BigData55660.2022.10020368>.
- Al-Kahla, W., Shatnawi, A.S. and Taqieddin, E. (2021) 'A Taxonomy of Web Security Vulnerabilities', *2021 12th International Conference on Information and Communication Systems (ICICS)*, pp. 424–429. Available at: <https://doi.org/10.1109/ICICS52457.2021.9464576>.
- Alouffi, B. *et al.* (2021) 'A Systematic Literature Review on Cloud Computing Security: Threats and Mitigation Strategies', *IEEE Access*, 9, pp. 57792–57807. Available at: <https://doi.org/10.1109/ACCESS.2021.3073203>.
- Aloufi, H.A. and Abdulaziz, A.H. (2022) 'Challenges and Obstacles Facing Data in the Big Data Environment', *Communications in Mathematics and Applications*, 13(1), pp. 331–349. Available at: <https://doi.org/10.26713/cma.v13i1.1974>.
- Alsmadi, I. (2019) 'Cyber Threat Analysis', *The NICE Cyber Security Framework* [Preprint]. Available at: [https://doi.org/10.1007/978-3-030-02360-7\\_9](https://doi.org/10.1007/978-3-030-02360-7_9).
- Alwaheidi, M.K.S. and Islam, S. (2022) 'Data-Driven Threat Analysis for Ensuring Security in Cloud Enabled Systems', *Sensors*, 22(15), p. 5726. Available at: <https://doi.org/10.3390/s22155726>.
- Alwaheidi, M.K.S., Islam, S. and Papastergiou, S. (2022) 'A Conceptual Model for Data-Driven Threat Analysis for Enhancing Cyber Security', in *International Conference on Interactive Collaborative Robotics*, pp. 365–374. Available at: [https://doi.org/10.1007/978-3-031-14054-9\\_34](https://doi.org/10.1007/978-3-031-14054-9_34).
- Amit, R. and Schoemaker, P.J.H. (1993) 'Strategic assets and organizational rent: Strategic Assets', *Strategic Management Journal*, 14(1), pp. 33–46. Available at: <https://doi.org/10.1002/smj.4250140105>.
- Armstrong, R. *et al.* (2011) 'Cochrane Update. "Scoping the scope" of a cochrane review.', *Journal of public health*, 33 1, pp. 147–50. Available at: <https://doi.org/10.1093/pubmed/fdr015>.
- Baker, W. *et al.* (2011) '2011 Data Breach Investigations Report'.

- Bala Bharathi, B. and Suresh Babu, E. (2018) 'A Novel Approach to Cyber Hazard Management Intelligence System', *International Journal of Engineering & Technology*, 7(2.7), p. 473. Available at: <https://doi.org/10.14419/ijet.v7i2.7.10866>.
- Barnum, S. (2014) *About STIX | STIX Project Documentation*. Available at: <https://stixproject.github.io/getting-started/whitepaper/> (Accessed: 16 September 2023).
- Bencsáth, B. *et al.* (2012) 'The Cousins of Stuxnet: Duqu, Flame, and Gauss', *Future Internet*, 4(4), pp. 971–1003. Available at: <https://doi.org/10.3390/fi4040971>.
- Berecki, B. (2019) 'How to Protect Data in Motion', *Endpoint Protector Blog*, 30 October. Available at: <https://www.endpointprotector.com/blog/how-to-protect-data-in-motion> (Accessed: 18 September 2023).
- Bonabeau, E. (2002) 'Agent-based modeling: Methods and techniques for simulating human systems', *Proceedings of the National Academy of Sciences*, 99(suppl\_3), pp. 7280–7287. Available at: <https://doi.org/10.1073/pnas.082080899>.
- Breed, D.G. and Verster, T. (2019) 'An empirical investigation of alternative semi-supervised segmentation methodologies', *South African Journal of Science*, 115(3/4). Available at: <https://doi.org/10.17159/sajs.2019/5359>.
- CIS Controls* (2023) *CIS*. Available at: <https://www.cisecurity.org/controls/> (Accessed: 17 September 2023).
- Cisco, N. (2015) *Top 10 Cloud Risks*. Available at: <https://docplayer.net/2298521-Top-10-cloud-risks-that-will-keep-you-awake-at-night.html> (Accessed: 19 September 2023).
- Collier, Z.A. *et al.* (2014) 'Cybersecurity Standards: Managing Risk and Creating Resilience', *Computer*, 47(9), pp. 70–76. Available at: <https://doi.org/10.1109/MC.2013.448>.
- Dekker, M. and Alevizos, L. (2024) 'A Threat-Intelligence Driven Methodology to Incorporate Uncertainty in Cyber Risk Analysis and Enhance Decision Making', *SECURITY AND PRIVACY*, 7(1), p. e333. Available at: <https://doi.org/10.1002/spy2.333>.
- Deshpande, I. (2020) 'What Is Customer Data? Definition, Types, Collection, Validation and Analysis', *Spiceworks*, 16 March. Available at: <https://www.spiceworks.com/marketing/customer-data/articles/what-is-customer-data/> (Accessed: 16 September 2023).
- Elahi, H. *et al.* (2021) 'On the Characterization and Risk Assessment of AI-Powered Mobile Cloud Applications', *Computer Standards & Interfaces*, 78, p. 103538. Available at: <https://doi.org/10.1016/j.csi.2021.103538>.
- European Commission MITIGATE Project* (2020) *CORDIS | European Commission*. Available at: <https://cordis.europa.eu/project/id/653212> (Accessed: 18 September 2023).
- Farboodi, M. *et al.* (2022) 'Valuing Financial Data'.
- Fortinet, N. (2021) *Advanced Security for SAP Solutions on Google Cloud*.
- Ganiga, R. *et al.* (2020) 'Security framework for cloud based electronic health record (EHR) system', *International Journal of Electrical and Computer Engineering (IJECE)*, 10(1), p. 455. Available at: <https://doi.org/10.11591/ijece.v10i1.pp455-466>.
- Giannoutakis, K.M. *et al.* (2020) 'Next Generation Cloud Architectures', in T. Lynn *et al.* (eds) *The Cloud-to-Thing Continuum*. Cham: Springer International Publishing (Palgrave Studies in Digital Business & Enabling Technologies), pp. 23–39. Available at: [https://doi.org/10.1007/978-3-030-41110-7\\_2](https://doi.org/10.1007/978-3-030-41110-7_2).

- Green, B.N., Johnson, C.D. and Adams, A. (2006) 'Writing narrative literature reviews for peer-reviewed journals: secrets of the trade', *Journal of Chiropractic Medicine*, 5(3), pp. 101–117. Available at: [https://doi.org/10.1016/S0899-3467\(07\)60142-6](https://doi.org/10.1016/S0899-3467(07)60142-6).
- Guarino, N. and Welty, C. (2009) 'An Overview of OntoClean', in, pp. 201–220. Available at: [https://doi.org/10.1007/978-3-540-92673-3\\_9](https://doi.org/10.1007/978-3-540-92673-3_9).
- Harris, M. et al. (2019) *Infrastructure Security and Segmentation > The Three Planes* | Cisco Press. Available at: <https://www.ciscopress.com/articles/article.asp?p=2928193> (Accessed: 16 September 2023).
- Hernan, S. et al. (2006) 'Threat modeling-uncover security design flaws using the stride approach', *MSDN Magazine*, pp. 68–75.
- Husák, M. et al. (2019) 'Survey of Attack Projection, Prediction, and Forecasting in Cyber Security', *IEEE Communications Surveys & Tutorials*, 21(1), pp. 640–660. Available at: <https://doi.org/10.1109/COMST.2018.2871866>.
- Hytönen, E., Trent, A. and Ruoslahti, H. (2022) 'Societal Impacts of Cyber Security in Academic Literature – Systematic Literature Review', *European Conference on Cyber Warfare and Security* [Preprint]. Available at: <https://doi.org/10.34190/eccws.21.1.288>.
- IBM Security X-Force Threat Intelligence Index 2023* (2023). Available at: <https://www.ibm.com/reports/threat-intelligence> (Accessed: 18 September 2023).
- Impe, K.V. (2018) 'What Are the Different Types of Cyberthreat Intelligence?', *Security Intelligence*, 4 June. Available at: <https://securityintelligence.com/what-are-the-different-types-of-cyberthreat-intelligence/> (Accessed: 16 September 2023).
- Insua, D. et al. (2019) 'An Adversarial Risk Analysis Framework for Cybersecurity', *Risk Analysis*, 41, pp. 16–36. Available at: <https://doi.org/10.1111/risa.13331>.
- ISO/IEC 27001, The Information Security (ISMS) Standard* (2022) <https://www.isms.online/>. Available at: <https://www.isms.online/iso-27001/> (Accessed: 17 September 2023).
- Ive, J. (2022) 'Leveraging the potential of synthetic text for AI in mental healthcare', *Frontiers in Digital Health*, 4. Available at: <https://doi.org/10.3389/fdgth.2022.1010202>.
- Jiang, Y. et al. (2023) 'Model-Based Cybersecurity Analysis', *Business & Information Systems Engineering*, 65(6), pp. 643–676. Available at: <https://doi.org/10.1007/s12599-023-00811-0>.
- Jodka, S. (2018) *The GDPR Covers Employee/HR Data and It's Tricky, Tricky (Tricky) Tricky: What HR Needs to Know* | Insights | Dickinson Wright. Available at: <https://www.dickinson-wright.com/news-alerts/the-gdpr-covers-employee-hr-data-and-tricky> (Accessed: 16 September 2023).
- K, B., T, Sudalaimuthu. and V, S.Solomi. (2023) 'An Analysis of Various Cyber Threat Modeling', in *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)*. *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, pp. 426–429. Available at: <https://doi.org/10.1109/ICAIS56108.2023.10073771>.
- Kaja, N., Shaout, A. and Ma, D. (2019) 'Fuzzy Based Threat Assessment Model (FTAM)', *2019 International Arab Conference on Information Technology (ACIT)*, pp. 144–149. Available at: <https://doi.org/10.1109/ACIT47987.2019.8991129>.

Kawanishi, Y. *et al.* (2023) ‘A Study on Threat Analysis and Risk Assessment Based on the “Asset Container” Method and CWSS’, *IEEE Access*, 11, pp. 18148–18156. Available at: <https://doi.org/10.1109/ACCESS.2023.3246497>.

Khan, N.A. and Manzoor Rashid, A.Z.M. (2022) ‘Action and Evidence-Based Research’, in M.R. Islam, N.A. Khan, and R. Baikady (eds) *Principles of Social Research Methodology*. Singapore: Springer Nature, pp. 279–290. Available at: [https://doi.org/10.1007/978-981-19-5441-2\\_19](https://doi.org/10.1007/978-981-19-5441-2_19).

Kirvan, P. (2023) *Top 10 threat modeling tools, plus features to look for | TechTarget, Security*. Available at: <https://www.techtarget.com/searchsecurity/tip/Top-threat-modeling-tools-plus-features-to-look-for> (Accessed: 5 October 2023).

Lakhanpal, A. (2020) ‘The Strategy and Planning in Cloud Computing in Business Planning’, 7(4).

Lee, B. *et al.* (2022) ‘Designing trustworthy IoT systems: Critical challenges and approaches for generating value’, *Human Interaction and Emerging Technologies (IHET 2022): Artificial Intelligence and Future Applications* [Preprint]. Available at: <https://doi.org/10.54941/ahfe1002783>.

Linnenluecke, M., Marrone, M. and Singh, A.K. (2020) ‘Conducting systematic literature reviews and bibliometric analyses’, *Australian Journal of Management*, 45, pp. 175–194. Available at: <https://doi.org/10.1177/0312896219877678>.

Liu, Y. *et al.* (2023) ‘Threat analysis and risk assessment for applying cybersecurity engineering’, in *International Conference on Electronic Information Engineering and Computer Science (EIECS 2022)*. *International Conference on Electronic Information Engineering and Computer Science (EIECS 2022)*, SPIE, pp. 324–332. Available at: <https://doi.org/10.1117/12.2668050>.

Luo, F. *et al.* (2021) ‘Threat Analysis and Risk Assessment for Connected Vehicles: A Survey’, *Secur. Commun. Networks*, 2021, pp. 12638201–126382019. Available at: <https://doi.org/10.1155/2021/1263820>.

Mauri, L. and Damiani, E. (2022) ‘Modeling Threats to AI-ML Systems Using STRIDE’, *Sensors*, 22(17), p. 6662. Available at: <https://doi.org/10.3390/s22176662>.

McNulty, A.A.F., Keith (2022) ‘Using Data in Organizations’, in *Data, Methods and Theory in the Organizational Sciences*. Routledge.

*Microsoft Threat Modeling Tool overview - Azure* (2022). Available at: <https://learn.microsoft.com/en-us/azure/security/develop/threat-modeling-tool> (Accessed: 30 September 2023).

*MITRE CAPEC* (2023). Available at: <https://capec.mitre.org/data/index.html> (Accessed: 18 September 2023).

*MITRE CWE* (2023) *CWE - CWE List Version 4.12*. Available at: <https://cwe.mitre.org/data/index.html> (Accessed: 18 September 2023).

Muscio, D. and Wilson, G. (2017) ‘Cyber Assurance: How Internal Audit, Compliance and Information Technology Can Fight the Good Fight Together’.

Newaz, A.I. *et al.* (2022) ‘Systematic Threat Analysis of Modern Unified Healthcare Communication Systems’, in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*. *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Rio de Janeiro, Brazil: IEEE, pp. 1404–1410. Available at: <https://doi.org/10.1109/GLOBECOM48099.2022.10001605>.

*NIST CYBERSECURITY FRAMEWORK* (2018) *NIST*. Available at: <https://www.nist.gov/cyberframework/getting-started> (Accessed: 17 September 2023).

Ntsiepdjap, B.S. (2022) ‘DYNAMIC RISK ASSESSMENT FOR CRITICAL INFRASTRUCTURES UNDER ATTACK’, *International Journal of Advanced Research* [Preprint]. Available at: <https://doi.org/10.21474/ijar01/15433>.

*OWASP Threat Dragon* | OWASP Foundation (2023). Available at: <https://owasp.org/www-project-threat-dragon/> (Accessed: 30 September 2023).

PCI Security Standards Council (2023) *PCI Security Standards Council*. Available at: [https://www.pcisecuritystandards.org/about\\_us/policies/](https://www.pcisecuritystandards.org/about_us/policies/) (Accessed: 17 September 2023).

Pires, S. and Mascarenhas, C. (2023) ‘Cyber Threat Analysis Using Pearson and Spearman Correlation Via Exploratory Data Analysis’, in *2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC)*. *2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pp. 257–262. Available at: <https://doi.org/10.1109/ICSCCC58608.2023.10176973>.

Portalatin, M. *et al.* (2021) ‘Data Analytics for Cyber Risk Analysis Utilizing Cyber Incident Datasets’, *2021 Systems and Information Engineering Design Symposium (SIEDS)*, pp. 1–6. Available at: <https://doi.org/10.1109/SIEDS52267.2021.9483743>.

Powell, M. (2019) *11 Eye Opening Cyber Security Statistics for 2019*, *CPO Magazine*. Available at: <https://www.cpomagazine.com/tech/11-eye-opening-cyber-security-statistics-for-2019/> (Accessed: 16 September 2023).

Pracht, D., Toelle, A. and Broaddus, B. (2022) ‘Action Research: A Methodology for Organizational Change: 4H424, 2/2022’, *EDIS*, 2022(1). Available at: <https://doi.org/10.32473/edis-4h424-2022>.

Robinson, S. (2008) ‘Conceptual modelling for simulation Part II: A framework for conceptual modelling’, *Journal of the Operational Research Society*, 59, pp. 291–304. Available at: <https://doi.org/10.1057/palgrave.jors.2602369>.

Sabillon, R. *et al.* (2017) ‘A Comprehensive Cybersecurity Audit Model to Improve Cybersecurity Assurance: The CyberSecurity Audit Model (CSAM)’, in *2017 International Conference on Information Systems and Computer Science (INCISCOS)*. *2017 International Conference on Information Systems and Computer Science (INCISCOS)*, Quito: IEEE, pp. 253–259. Available at: <https://doi.org/10.1109/INCISCOS.2017.20>.

Sadique, F. *et al.* (2018) ‘Automated Structured Threat Information Expression (STIX) Document Generation with Privacy Preservation’, in *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. *2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York City, NY, USA: IEEE, pp. 847–853. Available at: <https://doi.org/10.1109/UEMCON.2018.8796822>.

Saini, V., Duan, Q. and Paruchuri, V. (2008) ‘Threat Modeling Using Attack Trees’, *Journal of Computing Sciences in Colleges*, 23.

*SAP S/4HANA and Business Process Automation* (2023) *SAP*. Available at: <https://www.sap.com/mena/products/erp/s4hana.html> (Accessed: 19 September 2023).

Schmittner, C. *et al.* (2019) ‘Preliminary Considerations for a Cooperative Intelligent Transport System Cybersecurity Reference Architecture’, *2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE)*, pp. 1–6. Available at: <https://doi.org/10.1109/ICCVE45908.2019.8965116>.

Sharma, V., Poulouse, J. and Maheshkar, C. (2023) ‘Analytics Enabled Decision Making “Tracing the Journey from Data to Decisions”’, in V. Sharma, C. Maheshkar, and J. Poulouse (eds) *Analytics Enabled Decision Making*. Singapore: Springer Nature, pp. 1–22. Available at: [https://doi.org/10.1007/978-981-19-9658-0\\_1](https://doi.org/10.1007/978-981-19-9658-0_1).

Shen, W. *et al.* (2022) ‘Power Internet Assets Security Threat Assessment based on the Cost of Security Protection’, in *2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*. *2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*, pp. 171–174. Available at: <https://doi.org/10.1109/AEMCSE55572.2022.00042>.

Shevchenko, N. *et al.* (2018) ‘THREAT MODELING: A SUMMARY OF AVAILABLE METHODS’.

Siddaway, A.P., Wood, A. and Hedges, L. (2019) ‘How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses.’, *Annual review of psychology*, 70, pp. 747–770. Available at: <https://doi.org/10.1146/annurev-psych-010418-102803>.

Snyder, H. (2019) ‘Literature review as a research methodology: An overview and guidelines’, *Journal of Business Research* [Preprint]. Available at: <https://doi.org/10.1016/J.JBUSRES.2019.07.039>.

Sobers, R. (2022) *166 Cybersecurity Statistics and Trends [updated 2022]*. Available at: <https://www.varonis.com/blog/cybersecurity-statistics> (Accessed: 16 September 2023).

Souppaya, M. and Scarfone, K. (2016) *Guide to Data-Centric System Threat Modeling*. NIST Special Publication (SP) 800-154 (Draft). National Institute of Standards and Technology. Available at: <https://csrc.nist.gov/pubs/sp/800/154/ipd> (Accessed: 17 September 2023).

Splunk (2021) ‘why-you-need-improved-operational-intelligence-for-big-data.pdf’. Available at: <https://www.splunk.com/pdfs/why-you-need-improved-operational-intelligence-for-big-data.pdf> (Accessed: 3 September 2022).

Staron, M. (2020) ‘Action Research as Research Methodology in Software Engineering’, in M. Staron (ed.) *Action Research in Software Engineering: Theory and Applications*. Cham: Springer International Publishing, pp. 15–36. Available at: [https://doi.org/10.1007/978-3-030-32610-4\\_2](https://doi.org/10.1007/978-3-030-32610-4_2).

Straub, J. (2020) ‘Modeling Attack, Defense and Threat Trees and the Cyber Kill Chain, ATT&CK and STRIDE Frameworks as Blackboard Architecture Networks’, in *2020 IEEE International Conference on Smart Cloud (SmartCloud)*. *2020 IEEE International Conference on Smart Cloud (SmartCloud)*, Washington DC, WA, USA: IEEE, pp. 148–153. Available at: <https://doi.org/10.1109/SmartCloud49737.2020.00035>.

Tatam, M. *et al.* (2021) ‘A review of threat modelling approaches for APT-style attacks’, *Heliyon*, 7(1), p. e05969. Available at: <https://doi.org/10.1016/j.heliyon.2021.e05969>.

Taylor, T., Araujo, F. and Shu, X. (2020) ‘Towards an Open Format for Scalable System Telemetry’, *2020 IEEE International Conference on Big Data (Big Data)*, pp. 1031–1040. Available at: <https://doi.org/10.1109/BigData50022.2020.9378294>.

Teoh, C.S. and Mahmood, A.K. (2017) ‘National cyber security strategies for digital economy’, in *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*. *2017 5th International Conference on Research and Innovation in Information Systems (ICRIIS)*, Langkawi, Malaysia: IEEE, pp. 1–6. Available at: <https://doi.org/10.1109/ICRIIS.2017.8002519>.

*Threat Modeling Platform* (2023). Available at: <https://www.iriusrisk.com/threat-modeling-platform> (Accessed: 30 September 2023).

*Top Strategic Cybersecurity Trends for 2023* (2023) *Gartner*. Available at: <https://www.gartner.com/en/articles/top-strategic-cybersecurity-trends-for-2023> (Accessed: 17 September 2023).

Tounsi, W. (2019) ‘What is Cyber Threat Intelligence and How is it Evolving?’, in W. Tounsi (ed.) *Cyber-Vigilance and Digital Trust*. 1st edn. Wiley, pp. 1–49. Available at: <https://doi.org/10.1002/9781119618393.ch1>.

- Turner, A., McCombie, S. and Uhlmann, A.J. (2019) 'A target-centric intelligence approach to WannaCry 2.0', *Journal of Money Laundering Control* [Preprint]. Available at: <https://doi.org/10.1108/jmlc-01-2019-0005>.
- Ucedavélez, T. and Morana, M.M. (2015) *Risk Centric Threat Modeling: Process for Attack Simulation and Threat Analysis*. 1st edn. Wiley. Available at: <https://doi.org/10.1002/9781118988374>.
- Ullah, F. *et al.* (2018) 'Data exfiltration: A review of external attack vectors and countermeasures', *Journal of Network and Computer Applications*, 101, pp. 18–54. Available at: <https://doi.org/10.1016/j.jnca.2017.10.016>.
- Viswanathan, G. and Prabhu, J. (2021) 'A hybrid threat model for system-centric and attack-centric for effective security design in SDLC', *Web Intell.*, 19, pp. 1–11. Available at: <https://doi.org/10.3233/web-210452>.
- Wolf, M. and Serpanos, D. (2019) 'Threats and Threat Analysis', *Safe and Secure Cyber-Physical Systems and Internet-of-Things Systems* [Preprint]. Available at: [https://doi.org/10.1007/978-3-030-25808-5\\_3](https://doi.org/10.1007/978-3-030-25808-5_3).
- Wood, C. (2019) *Data-centric security should be the new focus of cybersecurity investments*. Available at: <https://www.linkedin.com/pulse/data-centric-security-should-new-focus-cybersecurity-investments> (Accessed: 16 September 2023).
- Zhang, H. *et al.* (2022) 'Cybersecurity Threat Assessment Integrating Qualitative Differential and Evolutionary Games', *IEEE Transactions on Network and Service Management*, 19(3), pp. 3425–3437. Available at: <https://doi.org/10.1109/TNSM.2022.3166348>.
- Zou, Q. *et al.* (2020) 'Automatic Recognition of Advanced Persistent Threat Tactics for Enterprise Security', *Proceedings of the Sixth International Workshop on Security and Privacy Analytics* [Preprint]. Available at: <https://doi.org/10.1145/3375708.3380314>.
- Zulkarnain, N. *et al.* (2021) 'Big Data in Business and Ethical Challenges', in *2021 International Conference on Information Management and Technology (ICIMTech)*. *2021 International Conference on Information Management and Technology (ICIMTech)*, pp. 298–303. Available at: <https://doi.org/10.1109/ICIMTech53080.2021.9534963>.