A contribution to Industry 4.0: a framework to secure standardised data exchange

Marion Toussaint

UNIVERSITÉ DE LORRAINE

École Doctorale IAEM Lorraine

THÈSE

Intitulée

Une contribution à l'industrie 4.0 : un cadre pour sécuriser l'échange de données standardisées

A contribution to Industry 4.0: a framework to secure standardised data exchange

Présentée et soutenue publiquement le 12 septembre 2022

pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

En Automatique, Traitement du signal et des Images, et Génie Informatique

par

Marion TOUSSAINT

Composition du jury

Président :

Prof. William Dérigent, Professeur à l'Université de Lorraine

Rapporteurs :

Dr. HDR. Olivia Penas, Ingénieure de Recherche, HDR**,** ISAE-SUPMECA

Prof. Ricardo Jardim-Goncalves, Professeur à l'université Nouvelle de Lisbonne, Portugal

Examinateurs :

Dr. Raphaël Barbau, Chercheur sénior au NIST, USA

Prof. Hervé Panetto, Professeur à l'Université de Lorraine (Directeur de thèse)

Dr. Sylvère Krima, Chercheur sénior au NIST, USA (Co-Directeur de thèse)

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisor Dr. Sylvère Krima for the continuous support, guidance and motivation. Thank you for believing in me.

I would like to thank my supervisor Prof. Hervé Panetto for sharing his experience and knowledge during my Ph.D.

Besides my supervisors, I would like to thank Prof. Ricardo Jardim-Goncalves and Dr. Olivia Penas, for accepting to be the official reviewers of this thesis, and the other members of the jury, Prof. William Derigent and Dr. Raphaël Barbau for examining my work.

I am very grateful to the National Institute of Standards and Technology (NIST) for sponsoring this work. Without NIST, this work would not have been performed. I would like to thank Allison Barnard-Feeney for the opportunity as well as the support, enthusiasm, and knowledge. Without your help, I would have not been able to fully understand and immersed in the standard development process. I would also like to thank all the members of the ISO experts that I had the chance to meet and that shared their experience and knowledge with me. Moreover, I would like to thank my colleagues at NIST, especially Robert Lipman and Edward Du.

In addition, I would like to thank my friends: Victoria, Constance, Yandé, Sarra, Alex-Kevin, Anne-Gaëlle, Israël, Laëtitia, Samia, Coline, Aimée, Mathilde, Maud, and many others for the support as well as their patience when I was talking about my thesis.

Finally, I express my profound love and gratitude to my parents, my stepmother, my brother, and sister. This achievement would not have been possible without you. Thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AGESIC	Agency Electronic Government and Information Society and Knowledge
AI	Artificial Intelligence
AID	Asset Identification
AM	Additive Manufacturing
ANSI	American National Standards Institute
AP 203	Application Protocol 203 – Configuration controlled 3D designs of mechanical parts and assemblies
AP 209	Application Protocol 209 – Composite and metallic structural analysis and related design
AP 210	Application Protocol 210 – Electronic assembly, interconnect, and packaging design
AP 214	Application Protocol 214 – Core data for automotive mechanical design processes
AP 239	Application Protocol 239 – Product life cycle support
AP 242	Application Protocol 242 – Managed model-based 3D engineering
AP	Application Protocol
API	Application Programming Interface
ARF	Asset Reporting Format
ART	Agile Release Trains
ASIC	Australian Securities and Investments Commission
BPM	Business Process Modeling
CAD	Computer-Aided Design
CCE	Common Configuration Enumeration
CCP	Connecting Corruption Propagation
CCPA	California Consumer Privacy Act
CCSS	Common Configuration Scoring System
CD	Continuous Deployment
CDP	Continuous Delivery Pipeline
CE	Continuous Exploration
CER	Critical Element Ratio
CI	Continuous Integration
CIA triad	Confidentiality, Integrity and Availability triad

CISO	Chief Information Security Officer
COBIT	Control Objectives for Information and related Technology
CPE	Common Platform Enumeration
CPS	Cyber-Physical Systems
CVE	Common Vulnerabilities and Exposures
CVS	Control Version System
CVSS	Common Vulnerability Scoring System
DAD	Disciplined Agile Delivery
DDoS	Distributed Denial-of-Service
DIS	Draft International Standard
EWIS	Electrical Wire Interconnect Systems
FDIS	Final Draft International Standard
GDPR	General Data Protection Regulation
HIPAA	Health Insurance Portability and Accountability Act
IACS	Industrial Automation and Control Systems
IBM	International Business Machines Corporation
ID	Identifier
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
ISA	International Society of Automation
ISMS	Information Security Management System
ISO	International Organization for Standardization
IT	Information Technology
JQL	Jira Query Language
LeSS	Large-scale Scrum
LGPD	General Personal Data Protection Law
MCU	Uruguayan Cybersecurity Framework
MITM	Man In The Middle
MTTC	Mean Time To Contain
MTTD	Mean Time To Detect

MTTI	Mean Time To Identify
MTTR	Mean Time To Respond
NIST	National Institute of Standard and Technology
NIST CSF	National Institute of Standard and Technologies Cybersecurity Framework
OCIL	Open Checklist Interactive Language
OMG	Object Management Group
OVAL	Open Vulnerability and Assessment Language
PCI DSS	Payment Card Industry Data Security Standards
PDAC	Plan, Do, Check, Adjust
PDM	Product Data Management
PI	Program Increment
PIPEDA	Personal Information Protection and Electronic Documents Act
PMI	Product Manufacturing Information
QIF	Quicken Interchange Format
RE	Requirement Engineering
ReqIF	Requirements Interchange Format
ROI	Return On Investment
RTE	Release Train Engineer
SAFe	Scaled Agile Framework
SAISO	Senior Agency Information Security Officer
SBA	Small Business Association
SC	Subcommittee
SCAP	Security Content Automation Protocol
SDO	Standard Development Organization
SOC	Security Operations Centre
STEP	Standard for the Exchange of Product model data
SWID	Software Identification
SysML	Systems Modeling Language
TC	Technical Committee
TMSAD	Trust Model for Security Automation Data
UML	Unified Modeling Language

URL	Uniform Resource Locator
XCCDF	Extensible Configuration Checklist Description Format
XMI	XML Metadata Interchange
XML	Extensible Markup Language
XP	Extreme Programming

SUMMARY

The recent digital transformation of the manufacturing world has resulted in numerous benefits, from higher quality products to enhanced productivity and shorter time to market. In this digital world, digital data has become a critical element in many critical decisions and processes within and across organisations. Data exchange is now a key process for the organisations' communication, collaboration, and efficiency. Industry 4.0 adoption of modern communication technologies has made this data available and shareable at a quicker rate than we can consume or track it. This speed brings significant challenges such as data interoperability and data traceability, two interdependent challenges that manufacturers face and must understand to adopt the best position to address them.

On one hand, data interoperability challenges delay faster innovation and collaboration. The growing volume of data exchange is associated with an increased number of heterogeneous systems that need to communicate with and understand each other. Information standards are a proven solution, yet their long and complex development process impedes them from keeping up with the fast-paced environment they need to support and provide interoperability for, slowing down their adoption. This thesis proposes a transition from predictive to adaptive project management with the use of Agile methods to shorten the development iterations and increase the delivery velocity, increasing standards adoption. While adaptive environments have shown to be a viable solution to align standards with the fast pace of industry innovation, most project requirements management solutions have not evolved to accommodate this change. This thesis also introduces a model to support better requirement elicitation during standards development with increased traceability and visibility.

On the other hand, data-driven decisions are exposed to the speed at which tampered data can propagate through organisations and corrupt these decisions. With the mean time to identify (MTTI) and mean time to contain (MTTC) such a threat already close to 300 days, the constant growth of data produced and exchanged will only push the MTTI and MTTC upwards. While digital signatures have already proven their use in identifying such corruption, there is still a need for formal data traceability framework to track data exchange across large and complex networks of organisations to identify and contain the propagation of corrupted data. This thesis analyses existing cybersecurity frameworks, their limitations, and introduces a new standard-based framework, in the form of an extended NIST CSF profile, to prepare against, mitigate, manage, and track data manipulation attacks. This framework is also accompanied with implementation guidance to facilitate its adoption and implementation by organisations of all sizes.

RÉSUMÉ

La récente transformation numérique de l'industrie a entraîné de nombreux avantages, allant de produits de meilleure qualité à une productivité accrue et à des délais de mise sur le marché plus courts. Dans ce monde numérique, les données numériques sont devenues un élément essentiel dans de nombreuses décisions et processus critiques au sein et entre les organisations. L'échange de données est désormais un processus clé pour la communication, la collaboration et l'efficacité des organisations. L'adoption par l'industrie 4.0 de technologies de communication modernes a rendu ces données disponibles et partageables à un rythme plus rapide que nous ne pouvons les consommer ou les suivre. Cette vitesse apporte des défis importants tels que l'interopérabilité des données et la traçabilité des données, deux défis interdépendants auxquels les industriels sont confrontés et doivent comprendre pour adopter la meilleure position pour les relever.

D'une part, les problèmes d'interopérabilité des données sont un frein à des innovations et des collaborations plus rapides. Le volume croissant d'échanges de données est associé à un nombre accru de systèmes hétérogènes qui ont besoin de communiquer et de se comprendre. Les normes d'information sont une solution prouvée, mais leur processus de développement long et complexe les empêche de suivre le rythme rapide de l'environnement qu'elles ont besoin de soutenir et d'assurer l'interopérabilité, ce qui ralentit leur adoption. Cette thèse propose une transition de la gestion de projet prédictive à adaptative avec l'utilisation de méthodes Agiles pour raccourcir les itérations de développement et augmenter la vitesse de livraison, augmentant ainsi l'adoption des normes. Alors que les environnements adaptatifs se sont révélés être une solution viable pour aligner les normes sur le rythme rapide de l'innovation de l'industrie, la plupart des solutions de gestion des exigences des projets n'ont pas évolué pour s'adapter à ce changement. Cette thèse introduit également un modèle pour soutenir une meilleure élicitation des exigences lors de l'élaboration des normes avec une traçabilité et une visibilité accrues.

D'autre part, les décisions basées sur les données sont exposées à la vitesse à laquelle les données falsifiées peuvent se propager dans les organisations et corrompre ces décisions. Avec un temps moyen pour identifier (MTTI) et un temps moyen pour contenir (MTTC) une telle menace déjà proche de 300 jours, la croissance constante des données produites et échangées ne fera qu'accroître le MTTI et le MTTC. Bien que les signatures numériques aient déjà prouvé leur utilité pour identifier une telle corruption, un cadre formel de traçabilité des données est toujours nécessaire pour suivre l'échange de données sur des réseaux vastes et complexes d'organisations afin d'identifier et de contenir la propagation des données corrompues. Cette thèse analyse les cadres de cybersécurité existants, leurs limites et introduit un nouveau cadre basé sur des normes, sous la forme d'un profil NIST CSF étendu, pour se préparer, atténuer, gérer et

suivre les attaques de manipulation de données. Ce cadre est également accompagné de conseils de mise en œuvre pour faciliter son adoption et sa mise en œuvre par les organisations de toutes tailles.

INTRODUCTION

Exchanging data entails ensuring that data remains understandable and accurate throughout the communication process. Data exchange is a complex process that includes many different problems. This thesis aims to describe and address two specific challenges related to the exchange of data. Section 1 provides a brief presentation of the industrial revolutions. Section 2 presents the main challenges related to data interoperability while Section 3 describes the main challenges related to data traceability. Finally, Section 4 discusses the approaches used to tackle these issues.

Problem setting

Over the centuries, technological advancement has changed the production methods that humans use. New techniques and production processes have radically changed people's working conditions and lifestyles.

The First Industrial Revolution marked the birth of mechanisation through the use of water and steam power. The Second Industrial Revolution reflected the emergence of mass production possible through the discovery of electricity. The Third marked the emergence of automation in production processes through the introduction of electronics and information technology. Finally, the Fourth Industrial Revolution, also known as “Industry 4.0”, was formed by the digital revolution based on Cyber-Physical Systems (CPS). It is also characterised by the interconnectivity of the systems and the access to real-time data.

The digital revolution in the world of manufacturing is fuelled by advances in information and communication technologies. Paper-based 2D drawings and unstructured data sources (e.g., spreadsheets, text documents, email, …) have been replaced by structured digital data models containing various types of information (e.g., product design, manufacturing equipment, process data …). On the same principle, automated processes to collect and analyse data in real-time have succeeded the manual methods formerly used.

The digitalisation of manufacturing, and the adoption of cyber-physical systems and IoT/CPS technology (e.g., smart sensors, smart actuators, machine learning) have facilitated and resulted in the generation of large volumes of heterogeneous data [1] (e.g., product models or telemetry data). Organisations produce, consume, and exchange massive volumes of data as part of their daily operations. Data now has the power to instantly turn into information, knowledge, and educated decisions, in an effort to boost performances (e.g., reducing cost or optimising resources) [2]. For instance, tooling data can now be processed and analysed by artificial intelligence (AI) agents to optimise machine performance and

energy efficiency in real-time. Digital data has become an essential player in many decision-making processes and a critical enabler in improving manufacturing competitiveness [3].

Industry 4.0 (or the so-called Industry of the Future in France) necessitates and enables fast access and exchange of that product data among a variety of applications and information systems - within and across organisations. A product creates and relies on a large amount of data during its lifecycle in response to different processes (e.g., design, manufacturing, inspection) and business needs (e.g., technical, commercial, regulatory). Every organisation involved in the product lifecycle relies on this data to perform its functions. It represents the "fuel" behind the organisation's contribution, efficiency, and value: organisations can create more value and drive faster innovation by exchanging data across them, facilitating collaboration.

Unfortunately, fast and reliable data exchange is also a complex operation that comes with multiple challenges [4], each of which can have drastic consequences on organisations, their operations, their products, and their collaborators. In this dissertation, we will focus on rising threats accompanying the rapid changes of the industry's dynamic environment, the security of data exchanges, and demonstrate how better **data interoperability** and **data traceability** can address them. In the next sections, we define and discuss the risks associated with two major and interdependent challenges, data interoperability and data traceability, inherent to the massive volume of data exchange caused by Industry 4.0 industrial digitisation.

Challenges related to data interoperability

Following this digital transformation of the industry and the modernisation of the adopted communication technologies, data is now available from all, to all, and in a multitude of formats. Organisations can easily connect different software and physical systems, internally and within their network of collaborators, as long as these systems speak a common language.

Unfortunately, today's manufacturing organisations are characterised by complex environments consisting of domain-specific components such as systems, networks, or machines, clustered in heterogeneous groups. While the interaction of these components is crucial for manufacturing as it supports production processes, effective interoperability across all elements of the product lifecycle is a growing challenge [5]. The amount of data produced and consumed continues to increase due to this growing ecosystem (of machines, systems, and networks), but so does the number of data formats. These data are collected from distributed data sources and therefore do not necessarily share the same format. Data heterogeneity is an important factor in data exchange. The different components of an organisation's environment must be able to unambiguously interpret, use, integrate and compare the information exchanged.

These different systems need a common language to exchange and understand information. The use of neutral model-based data standards helps provide a common data format, and thus facilitates technical interoperability between all parties involved in an exchange. Information standards are essential for properly integrating, exchanging, and interpreting data manufacturers rely on [6]. Standards define an agreed-upon language (data format, definitions, etc.) for data exchange between the different systems that consume, process, and produce data. The lack of standardisation results in a multiplication of information formats that are not necessarily compatible with each other, making it difficult for stakeholders to communicate and exchange data, as shown in Figure 1.

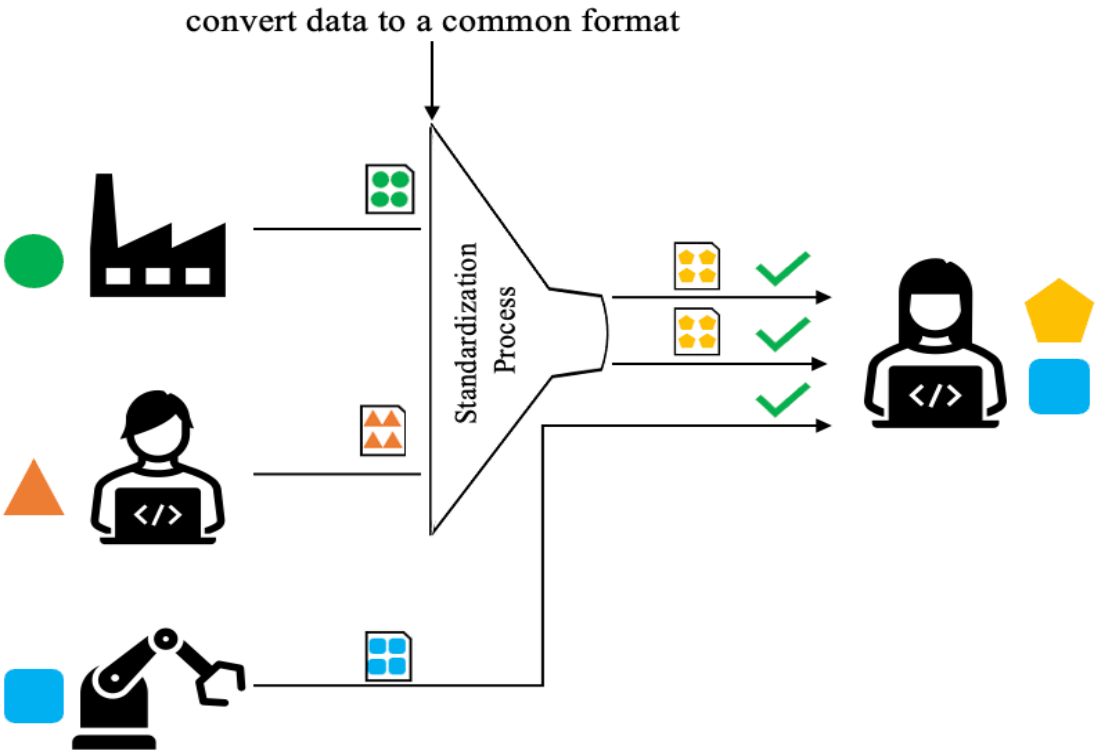

Figure 1 – Information standards provide a common language, enabling technical interoperability between systems.

Information standards are an important asset for organisations because they help facilitate business interaction and compatibility between systems, as well as support technical interoperability between systems, people, and organisations [7]. Information standardisation also saves time and reduces costs by eliminating the need to have separate translators for each pair of systems that need to exchange data. The adoption and implementation of standards by organisations improves performance, competitiveness, and transparency, given that standards promote the accessibility of information by all stakeholders. Information standards are powerful tools for productivity and innovation as they provide common language, vocabularies and platforms for organisations to collaborate and develop new technological solutions. Standards also support innovation by increasing the speed at which organisations bring new products to the

marketplace, thanks to faster and more accurate collaboration. Information standards are therefore key enablers to the evolution and digitalisation of the manufacturing sector. Nowadays, standards support the full product lifecycle. Product definition data is represented in ISO 10303 (informally known as STEP) [8]. Manufacturing planning systems can read in STEP data and generate manufacturing instructions in G-code [9] or ISO 10303-238 (STEP-NC) [10]. MTConnect Agents [11] stream machine execution data that represents an as-manufactured product. Coordinate measurement system software can read in STEP product definition data and generate inspection plans and inspection results represented in the Quality Information Framework (QIF) [12].

Despite this, information standards present a major challenge, which can impact their adoption and implementation by organisations: the complexity and current development process length of prominent standards are incompatible not only with the needs and pace of the industry but also with the lifespan of data.

The information standards development process is complex. This process is generally long, irregular, and difficult to plan. Firstly, the waterfall methodology for project management is prominent, which implies that 1) the entire deliverable is only available (for review) at the end of the development iteration, and 2) the requirements must be defined at the beginning of the project and do not change throughout the entire iteration. According to ISO itself, standard development iterations can last between 18 and 48 months [13]. This means that a new requirement identified after a new iteration just started, will not be addressed for another 18 to 48 months. But in both cases, additional time must be given to software vendors to implement, test, and deliver updated software solutions.

Secondly, standards are developed by experts that are working for different organisations. The contribution and participation of these experts to the standards development process are entirely voluntary. The resources available depend on the experts' schedules and their organisations' needs, making the development process irregular and difficult to plan.

The duration and management of the standards development process are not aligned with the needs of the industry. Strong market competition results in shortened product life cycles and requirements that change often and faster than the pace of standards development. The standards development process is incompatible with the data lifespan. Industry 4.0 values speed and rapid innovation. Consequently, manufacturers need standards development organisations to accelerate and simplify the standards development process, so the resulting standards represent current industry needs and are eagerly adopted.

Challenges related to data traceability

Manufacturing has become more automated, connected, and data-centric. Industry 4.0 is characterised by the networking of machines, systems, and products and the convergence of physical, digital, and virtual environments. This continuous networking and emergence of cyber-physical environments allow data to be more quickly accessible and facilitates the fast and timely exchange of data between the systems that require the information and the systems that have the information [14]. These data exchanges are both intra- and inter-organisational, and can be characterised as high-speed, high-volume, high-frequency, and low latency exchanges. For instance, on the manufacturing floors, complex instructions and monitoring data are exchanged in real-time between the different manufacturing systems. Similarly, the integration with other technologies such as artificial intelligence (AI), means that data is used to generate and share decisions at a pace and volume significantly greater than anything humans can manually validate or track.

This pace and volume of data exchange in the manufacturing world comes with significant challenges. The heavy reliance on data-driven decisions and the integration of new technologies have made organisations more vulnerable to cyber threats, a major concern for companies regardless of their size and sector. The manufacturing sector generates large amounts of data and relies heavily on it, which makes this sector an ideal target for cyber-attacks. To no surprise, the manufacturing sector was particularly impacted by cybercrime in 2020 and 2021. According to the IBM Security's X-ForceThreat Intelligence Index 2022 report, manufacturing was the most attacked sector in 2021 (with 23,2% of all attacks) while it was ranked second in 2020 (with 17,7% of all attacks) and eighth in 2019 (with 8,1% of all attacks) [15].

Generally, security threats are classified according to the governing principles of the CIA triad security model: confidentiality, integrity, and availability [16,17]. Data confidentiality requires that data remain secret or private, data integrity requires that data is trustworthy and free from tampering, and finally, data availability requires that data is always accessible to authorised access when it is needed. Threats and vulnerability are assessed based on the type of risks associated with and the potential damage they can cause to an organisation's assets such as data, applications, and systems.

These risks cannot always be averted and are a significant challenge to identify and contain. The IBM's Cost of a Data Breach Report 2022 shows that in 2021, the mean time to identify (MTTI) a data breach was 212 days and 75 days to be contained (MTTC) for a total lifecycle of 287 days. This represents a slight increase over 2020, when the average time to identify and contain a data breach was 280 days (an average of 207 days to identify and an average of 73 days to contain) [18].

One threat particularly relevant is data manipulation, an attack that focuses on subtly altering data [19] with the objective of manipulating data-driven decisions and that relies on data exchange to propagate tampered data and decisions across an organisation and its network. This tampering can result in corruption, modification, and/or destruction of the data, ultimately causing a loss of trust in the data [20] and the decisions derived from it. Data manipulation can also potentially lead to different manufactured products.

When data is exchanged, it leaves the private and trusted system of the data owner to be sent to other systems. This process presents the critical risk that the data exchanged might have been tampered with by unauthorised parties (see Figure 2). It is therefore important to ensure the data remains accurate, authentic and trustworthy during the entire exchange process. Data integrity is the CIA triad aspect the most impacted by data exchanges, as integrity assists both the sender, who must ensure that data attributed to them is not tampered with, and the receiver, who needs "the guarantee that the message that is sent is the same as the message received and that the message is not altered in transit" [21].

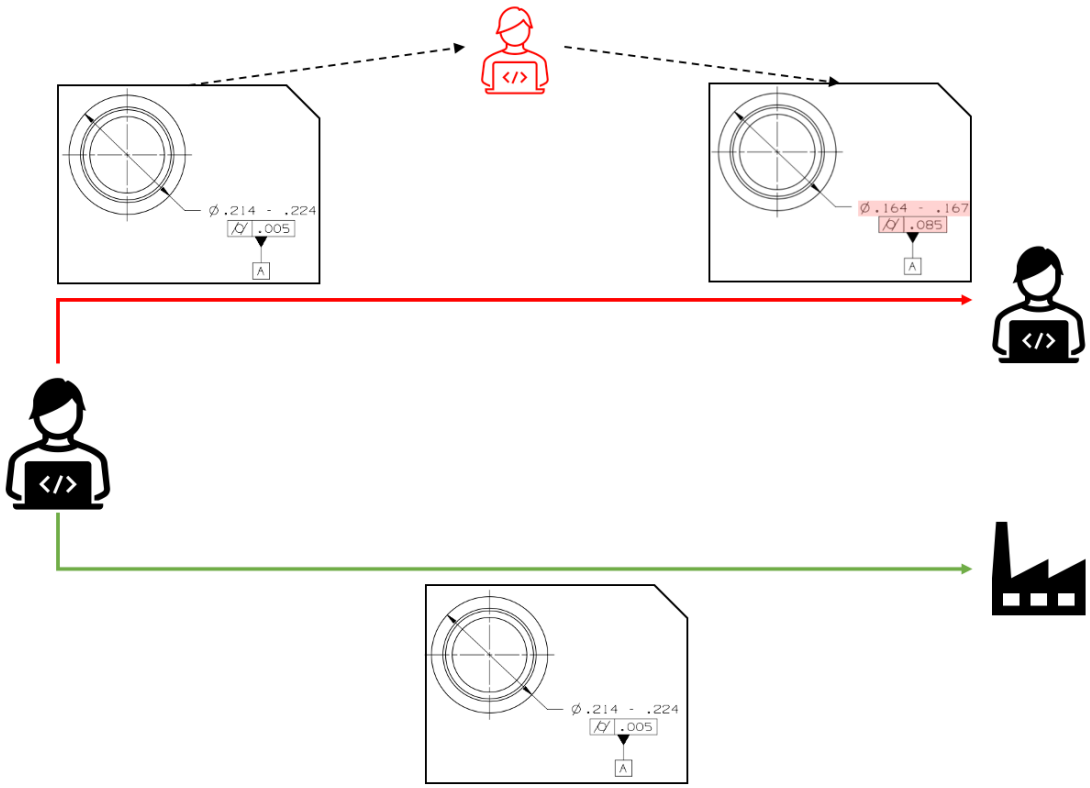

Figure 2 – Example of data manipulation during a data exchange – the red flow indicates a malicious actor tampering Product Manufacturing Information (PMI) on a 2D drawing using a Man In The Middle (MITM)[1] attack.

[1] NIST Computer Security Resource Center glossary [https://csrc.nist.gov/glossary/term/man_in_the_middle_attack]

Data integrity presents two main challenges: 1) validating the accuracy of the data and 2) tracking down inaccuracy. The former is commonly solved using digital signatures [22], while the latter is more complex and one that still needs to be addressed. The complexity (i.e., number of actors and steps involved), pace, and volume of data exchange that organisations are part of makes it impossible to manually account for and track down every single inaccuracy. Those same benefits and advantages that make manufacturers more competitive and innovative also make them more vulnerable to data integrity attacks.

Summary

As a result of the digital transformation of manufacturing, environments, and processes have become more connected, automated, and data-driven. In this digital world, data is now a key enabler in processes, exchanges, and decision-making. Data exchange has become a critical process for organisations' communication, collaboration, and efficiency. Manufacturing is heavily reliant on data and the exchange of this data between the different stakeholders, machines, and systems. The digitisation of Industry 4.0 with the emergence of new technologies and networked data sources supports new opportunities for organisational collaboration through high-speed and high-volume data exchange. However, the growing volume and speed at which data are exchanged bring significant challenges. In this thesis, we will address two of these challenges that manufacturers face, data interoperability and data traceability.

Data interoperability challenges have an impact on fast collaboration. The growing volume of data exchange is associated with an increased number of heterogeneous systems that need to communicate with and understand each other. Information standards are a recognised solution for supporting interoperability between systems and meaningful data sharing, as well as enhancing security. The digitisation of Industry 4.0 comes with technological innovations and fast changing market demands as well as rapidly changing and evolving cyber threats, and therefore standards must adapt very quickly. The standard development process is long and complex, which prevents the standards from keeping up with the fast-paced environment they need to support and provide interoperability for and therefore, slows down their publication and adoption.

Data traceability challenges have an impact on decision and data trustworthiness, a key element to achieving benefits such as supply chains optimization and manufacturing challenges. Leveraging new information and communication technologies has enabled a new data-driven approach to business decisions and processes, but it has also significantly increased the number and complexity of the associated cybersecurity challenges. In this context, data is exchanged across supply chains and distributed systems within and across organisations, which means a greater vulnerability surface for organisations and therefore an increased risk of attacks. Data is now exposed to numerous cyber threats, including data manipulation.

Data-driven decisions are vulnerable to the speed at which tampered data can propagate through organisations and corrupt them. There is a need for a formal data traceability solution to track data exchange across large and complex networks of organisations in order to identify and contain the propagation of corrupted data as quickly as possible. In this thesis, we will address data traceability issues by 1) demonstrating the security benefits of standard-based data exchanges, and 2) defining a solution to ensure the traceability of standard-based data exchange in order to limit the consequences data manipulation can have on organisations.

Digital data is a critical asset for organisations as it supports organisational decision-making and strategy. Organisations have to manage and exchange large volume of data while maintaining their integrity, quality and trustworthiness. It is critical to address data interoperability and data traceability challenges because data exchanges need to be understandable by all actors (stakeholders, machines, systems) and accurate in order to ensure that the decisions and processes based on data exchanged are trustworthy.

Structure of the document

The previous sections discussed the risks associated with two major challenges, data interoperability, and data traceability. In order to address these two challenges, this thesis focuses on the use of information standards and proposes two approaches: 1) to improve the traditional information standard development process to support the Industry 4.0 fast-paced environment and 2) to promote the use of standards and data traceability to help mitigate cyber risks in this complex environment.

Challenge 1: Data interoperability

To address the data interoperability issue, our goal is to facilitate the adoption of standards by improving the standards development process. Standards provide a common language that helps the communication and interpretation of information between systems.

Chapter 1 introduces information standards and their development lifecycle. This chapter also presents the limitations of this development process that slow down the adoption of standards.

Chapter 2 introduces the Agile methodology to address some of the standards development process limitations related to resource management. This chapter details a test case of using Agile in the context of standards development. It also presents the benefits and challenges that adopting Agile can bring to standards development.

Chapter 3 focuses on resource traceability, a main challenge of the standards development process. This chapter presents a review of existing challenges in requirements engineering. This chapter also proposes a Requirement Elicitation Model to address the requirements and resource traceability.

Challenge 2: Data traceability

In response to the cyber threats in manufacturing, our goal is to focus on data traceability and data manipulation risk management, that can help mitigate these risks.

Chapter 4 introduces the basic notions of cybersecurity and cybersecurity frameworks. This chapter also presents the cybersecurity challenges in the context of manufacturing.

Chapter 5 presents a review of cybersecurity frameworks with the intent to find one that could be leveraged for addressing data manipulation risks.

Chapter 6 proposes a Cybersecurity Framework Profile for data manipulation risk management. This chapter also presents data traceability, and defines implementation recommendations, including the definition of user stories and security metrics, to facilitate the adoption and implementation of the Profile.

Conclusion and future work

Finally, the last chapter presents the conclusions of this thesis and the perspective for future research works. This chapter discusses our contributions to address data interoperability and data traceability challenges. This chapter also proposes avenues to 1) extend and complement the Requirement Elicitation Profile we developed, and 2) to leverage the Data Manipulation Profile we presented through the development of a tool.

CHALLENGE 1: DATA INTEROPERABILITY

CHAPTER 1: INTRODUCTION TO STANDARDS DEVELOPMENT

This chapter introduces information standards and details their development lifecycle. This chapter also explains the complexity of this process and how it impacts the adoption of standards. Section 1.1 describes the standards development process. Section 1.2 presents ISO 10303: Automation systems and integration — Product data representation and exchange, a large and complex standard widely used in the manufacturing industry for the exchange of data. Section 1.3 presents the limitations related to the standards development process.

1.1 Standards development process

An information standard is a formal definition of how to represent and process domain-specific information. This definition is agreed upon by a community of experts, brought together within a Standard Development Organisation (SDO). An SDO's primary goals are to establish, promote, and maintain standards to meet the needs of users outside of that organisation. In other words, SDOs aim to create uniformity and consistency in terminology, product specifications, protocols, and more, across interested parties such as producers, consumers, government agencies. There are many SDOs across the world, among the best known are International Organization for Standardization (ISO), Object Management Group (OMG), Institute of Electrical and Electronics Engineers (IEEE) and American National Standards Institute (ANSI). Information standards are developed to ensure optimal interoperability and compatibility between information assets in order to exploit and process them in a consistent fashion. The use of standards offers practical application of expert knowledge and improves communication. Information from different providers can be represented, exchanged, and integrated together, seamlessly. Information standardisation can help facilitate business interaction through system integration, or support interoperability between technologies. Information standards are used in many domains such as manufacturing, healthcare, security, and more. Among the most well-known standards, we find ISO 9001 for quality management, ISO 14000 for environmental management, or ISO 27001 about information security management systems.

Standards are created in response to needs identified by domain/market stakeholders which can be from industry, government, or consumers. Figure 1.1 depicts the standard development process according to ISO. The standard development process begins at the proposal stage (10) when requirements are gathered to create a new work item proposal that is then submitted by these stakeholders, which we will call the Sponsor, to a Standards Development Organisation. The SDO is studying the request to determine if the creation of a standard on this particular subject will add value. This study entails extensive discussions to ensure that there is support for the proposed development of a standard on this subject and especially that

market stakeholders and other actors will allocate resources to the development of the standard. If the proposal is accepted, it officially becomes a draft standard and is assigned to a relevant technical body. This technical body will have been constituted through an open and consensus-based procedure with a specified scope and sector. Standards are developed by a system of committees and sub-committees and their associated working groups. These committees and sub-committees are further composed of "participating" members, who are people who have shown a desire to participate actively in the work of specific standards. The initial draft of a standard, called working draft, is normally done by a working group, which is appointed by the parent committee and is made up of a collaborative team of domain experts. The experts debate and agree on the elements they believe should be included in the standard. This work is done during the preparatory stage (20). Once they come to an agreement and the draft is finalised, the committee stage (30) begins during which the draft is reviewed, commented on and then voted on by the members of the parent committee. Once the draft is approved by the parent committee, the draft is submitted as a Draft International Standard (DIS) to all SDO members, during the enquiry stage (40) for them to comment on and vote on. If the draft is approved and no technical changes are made to the draft at stage 40, the process moves directly to the publication stage (60), the final text is published as an International/Official Standard. In the event that technical changes are requested during the inquiry stage, whether the DIS has been approved or not, the approval stage (50) becomes mandatory. The draft is modified following the comments at the previous stage, to incorporate the technical changes, then the new version of the text is submitted as the Final Draft International Standard (FDIS) to all SDO members. Once the FDIS is approved, the standard is published. The standard is then available to any entity that would be interested in using it, without any constraint, in processes, products, or services.

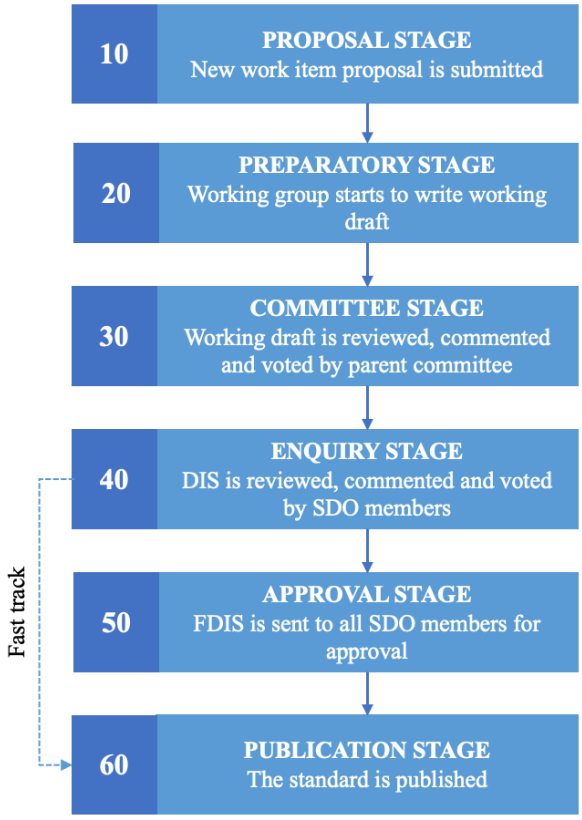

Figure 1.1 – Stages of the standard development process [23]

1.2 Introduction to ISO 10303 - Standard for the Exchange of Product data

ISO 10303, *Automation systems and integration – Product data representation and exchange* standard series, informally known as STEP, is one of the oldest and largest international information standards [24]. STEP is also the main standard for exchanging product data between manufacturing systems. For these reasons, STEP offers a great example of the standard information development process. We decided to use STEP as a reference to study this process and determine the limitations it may present.

In design, manufacturing, and product support processes, many systems are used to define technical product data. Each system has its own data format, so the exchange of information between each system of the lifecycle necessitates adapting this same information in the format of each system leading to redundancy and errors. Design data is complex, essential, and has to be shared throughout the product life-cycle, increasing the scope for errors and misunderstandings between systems. Data incompatibility in manufacturing extends costs and production time.

For these reasons, in 1984, a grand joint effort started under the International Organization for Standardization (ISO) have been made to produce a comprehensive standard for the electronic exchange of

product data between product life-cycle systems. At first, this standard was developed for design and manufacturing applications.

STEP aims to provide a complete and unambiguous description of manufacturing products, usable throughout their life cycle, regardless of the IT support used. The scope of STEP is much broader than other existing Computer-Aided Design (CAD) data exchange formats and it is intended to handle a wide range of product-related data covering the life-cycle of a product.

The development of STEP is one of the largest projects that ISO has ever carried out and six hundred people from many different countries have been involved for the last thirty-eight years. The project is still as active today as it was at its beginning. The development of ISO 10303 is the responsibility of the ISO/TC184/SC4 committee. STEP is increasingly recognised by the industry as an effective standard for exchanging product data between CAD systems. The standard was published for the first time in 1994 after ten years of work. Much of this project was to develop an infrastructure for the standard that will ensure its extensibility for the future.

STEP enables manufacturing applications to exchange and share data regardless of the specificities of the different systems that exchange and share information. ISO 10303 covers the whole product types (electronic, mechanical, architectural, furniture...) and all stages of the life-cycle (design, analysis, planning, manufacture, maintenance...). This range has expanded a lot since its creation because the standard is split into many parts that have been issued over the last three decades. Currently, the standard is composed of about 500 parts.

This standard is used to exchange data between different software in manufacturing. The support of this standard takes on a strategic aspect in the field of data exchange, in order to ensure that applications can communicate with each other in a collaborative framework.

Additionally, a product generates and uses various types of information from the beginning to the end of its life cycle. This information is created and exploited by various systems or software at different times and sometimes over several sites. So, to satisfy such constraints, it requires maintaining the integrity and consistency of the product data during exchanges. It is necessary that these data have a common form, interpretable by computers.

Therefore, STEP provides a means to describe the relevant information of a product, for all stages of its life cycle, relying on a set of general formalisms. STEP's objective is to enable communication between various systems of a company or between several companies, fulfilling these criteria: independence of hardware and software platforms, software interoperability, data longevity, scalability so it's possible to take into account the technical innovations without having to translate old data, modularity

to allow the exchange of the various types of data involved in the product life-cycle. ISO 10303 has three main principles that are: define an architecture for product data, support standardisation for manufacturing applications and define requirements for implementation of product data exchange.

The architecture of STEP consists of a set of standards whose objective is either the specification of a method of validation or the implementation of the definition of standard data schemas. To support its requirements, the STEP community developed a set of information modeling technologies: a library of reusable concepts to develop data models, ISO 10303-11, also known as the EXPRESS language [25], a file format to represent STEP data, ISO 10303-21 [26], an XML representation of STEP data, ISO 10303-28 [27] and other information models.

1.3 Limitations of the standards development process

The development of standards is a relatively long process that entails the collaboration of many people from various organisations. The members of a standard working group work on different projects for their organisations while contributing at the same time to the development of the standards. As a result, their contribution to the development of standards is voluntary. The resources available depend on the experts' schedules and the needs of their organisations, making the standards development process long, irregular, and difficult to plan. Furthermore, in some cases, the complexity of the standards rests on their architecture and application domains. The development of STEP is one of the largest initiatives that ISO has ever carried out. Six hundred people from many different countries have been involved for the last thirty-five years [24]. STEP is a product-centric standard that covers a wide range of industries, resulting in a large and multi-disciplinary community and requiring the collaboration of a large number of subject-matter experts. The STEP architecture had to evolve to support STEP development and growth, transitioning to a modular architecture “to enable the more efficient implementation and deployment of STEP standards without changing the fundamentals of the current technical architecture” [30], which means creating new modules by reusing, integrating, and extending existing Application Protocols (APs). Thus, APs are “more interoperable, easier to understand and manage, and quicker to develop” [31].

1.3.1 Waterfall methodology

The Waterfall model is a linear, sequential project management approach. Waterfall aims for predictability and does not encourage change. It emphasises a logical progression of phases. A project using the waterfall approach is divided into several phases. Each phase can only begin when the previous phase is completed. More generally, each phase of the project depends on the deliverables of the previous phase.

Generally, the waterfall methodology was divided into six distinct phases, as represented in Figure 1.2. However, with applications of this model in different areas, the number, and name of the phases may vary. The waterfall phases were originally defined as follows:

- Requirements: It is the initial phase of the project in which all the project requirements are gathered. The team collects as much information as possible to ensure the success of the project. This is a crucial phase because it gives the team all the information to plan the other phases of the project and once this phase is complete, requirements cannot be changed or added until the project is finished.
- Design: Once the requirements are specified, the team works on the overall product architecture and creates the product design.
- Implementation: Based on the requirements and the specifications from the previous phases, the team begins the full development process to build the final product.
- Testing: The pieces developed in the previous phase are tested and then integrated together. The entire product is tested for any faults and failures that need to be fixed before the deployment phase.
- Deployment: Once the testing is complete, the project is deployed, that is the final deliverable of the project is launched and delivered to the customers.
- Maintenance: Following the launch of the product, some issues can arise on the client-side. The team works on new versions of the product which once released will fix the problems as well as improve the product.

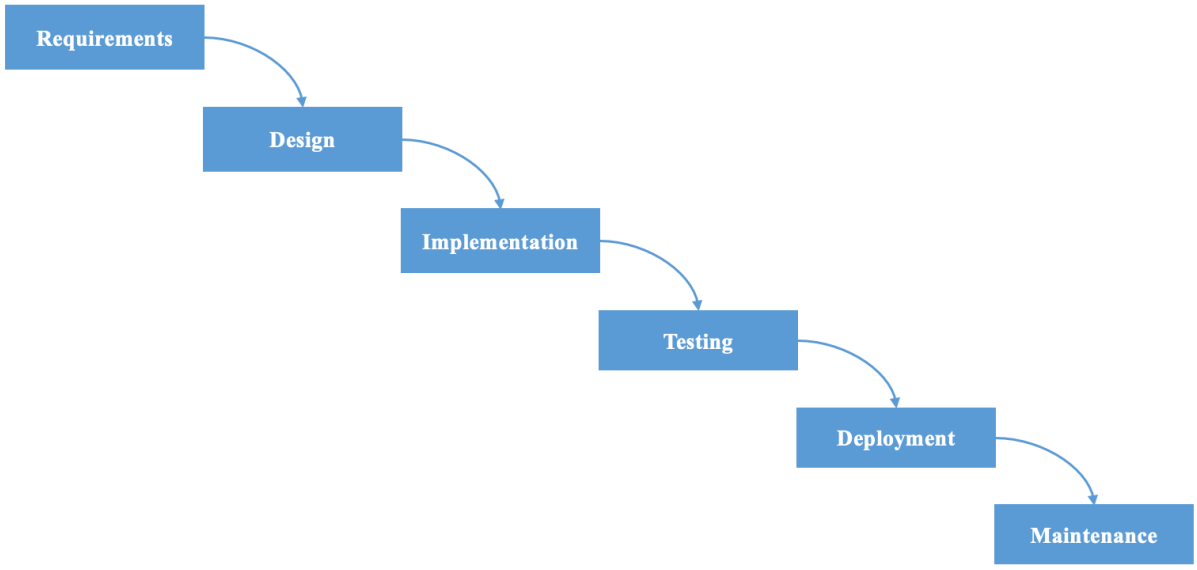

Figure 1.2 – Waterfall model

On one hand, the Waterfall model is simple to use and easy to understand and maintain. Each phase is clearly defined and has a specific objective. All processes and outcomes of each phase are well

documented. This model is ideal with clear requirements and goals. The phases of the project are already predefined; therefore, it is easy to measure the progress of the project. On the other hand, the waterfall model presents several challenges. First of all, it is extremely difficult to make changes in the previous phases. Similarly, it is not possible to change the requirements once the initial phase is complete, which does not allow for many revisions. Then, the testing process only begins once the implementation phase is over, which increases the risks, on one hand, of finding faults and defects later in the process and, on the other hand, that problems and challenges from earlier phases go unnoticed. Problems and defects found late in the development process will be expensive to fix. Finally, customers and other end-users no longer intervene in the project after the requirement phase until the end of the iteration.

1.3.2 Requirement management

As mentioned previously, the standards development process is lengthy. According to ISO, standards development iterations last between 18 and 48 months [13]. To add to this development time length, the most common project management method used when developing standards is the waterfall project management method. Until recently, ISO 10303 STEP standards were primarily developed using a traditional project management method (also known as predictive or waterfall approach).

In a waterfall project, the requirements are defined at the start of the development, during the requirements phase. Once this phase is completed, the requirements cannot be changed - or only in a rare instance and in a very careful and formal way - for the whole duration of the development iteration. The requirements identified afterward will have to wait until the end of the current iteration to be addressed.

As shown in Figure 1.3, in the case of standards development, this means that, if a new requirement R is identified during iteration 1 (but after the requirements phase):

- In the best-case scenario, iteration 1 lasts 18 months. This requirement will therefore be addressed and published after 18 months, at the end of iteration 2.
- In the worst-case scenario, iteration 1 lasts 48 months. As a result, this requirement will be addressed after 48 months and will be published up to 96 months later.

It is important to note that this requirement will be addressed in a published version of the standard, but it will not necessarily have been implemented.

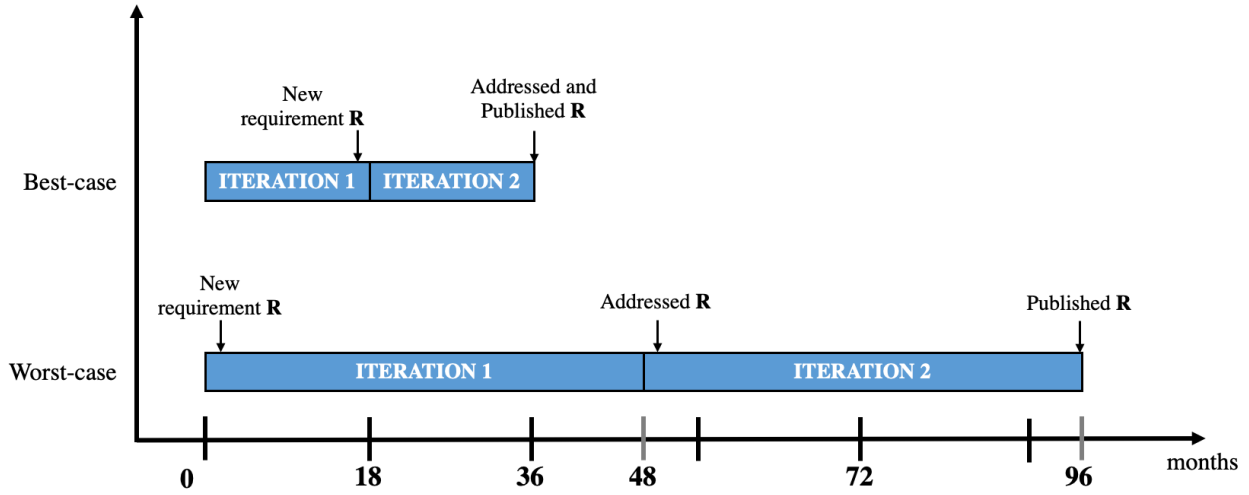

Figure 1.3 – Requirements management in standards development process

The fast-paced technological changes associated with Industry 4.0 question the traditional process of developing standards. The current development process is confronted with keeping pace with and supporting these rapid industrial developments. The current length of standards development iterations does not meet the needs of the industry for faster releases and incremental functionality [6]. In the manufacturing industry, strong market competition leads to a short product life cycle and as a result, the industry requirements change often and faster than the standard development. During these long development iterations, new technologies, processes, and needs are developed, so the standards team releases products and features, and uses technologies that are no longer necessarily relevant to industry and user realities. Standards risk missing opportunities to support the launching of new products and services, and at worst, they may hinder the effective launching of new innovations. The quality and time length of the standard development must be addressed in order to meet industry demand and maximise the benefits of standards.

Moreover, the standard development process involves numerous stakeholders from various organisations. As a result, requirements come from different sources and at various stages of the development cycle: each stakeholder has needs that must be met using the standard, and they can introduce new requirements at any moment during the development process. These new requirements are not always discussed during official meetings. Stakeholders may sometimes introduce requirements during informal discussions. As explained previously, it can take several months or even years to address the requirements and it is necessary to formalise and correctly document them in order to avoid losing any information during these long development iterations. The requirements should be structured and defined in a computer interpretable manner in order to correctly manage and validate accurately and consistently. The

formalisation of requirements would help ensure the traceability of the requirements from their source through their completion, thus, allowing the author of the requirement to confirm if it meets their needs.

1.3.3 Consensus management

Standards experts work for different organisations that are geographically dispersed around the world. Because of that, they cannot always be completely dedicated to the standard development and attend all the meetings and work sessions. It is critical to keep track of and document all the activities and decisions that are discussed and taken during the meetings. The current meeting minutes are not necessarily very detailed, formalised, and accessible. Then, it can be difficult to know exactly what happened during the meetings and stay up to date with the past and current activities of the development process.

In the case of STEP development, most teams used a system like Bugzilla for coordination after a new work item was approved or a bug was raised. Bugzilla was created as a general-purpose bug-tracking tool, but it has been expanded beyond that scope during the STEP development. It was used, until recently, as a requirements management, collaboration/consensus tracking, issue management, and task management system as well as for source model version control. The use of inadequate tooling like that caused a lot of confusion and maintenance, which extended development cycles and led to multiple rework cycles.

1.3.4 Human resource commitment and management

The standards development is primarily supported by the basis of volunteer resources from participating companies and organisations. The standards development is not a priority for most organisations and therefore it rests on the voluntary contribution of the experts, which leads to a lack of continuity and consistency. For example, the accessible number of resources fluent in STEP development is decreasing due to attrition and competing priorities [6]. Managing volunteer staff comes with some challenges:

- Currently, recruitment of new staff is informally done through national standards body nominations, which provides a limited pool of resources.
- New staff faces the difficult task of getting up to speed and bringing value to the standards development process (e.g., installing development software frameworks, completing training, and integrating into the project operating cadence to contribute).
- Developer resources are shared among the different standards development teams, which makes the demand for resources exceed the supply of volunteer time.

Furthermore, standards are developed by distributed teams because the members of standards development teams are geographically dispersed. Distributed teams help to overcome geographical barriers

while allowing more flexibility in the development process. However, this style of collaboration creates communication challenges such as time zone differences, cultural differences, and differences in communication style preference, which can affect the teams' efficiency.

Managing volunteer staff and working with distributed teams create challenges that can slow down the development of standards.

1.3.5 Integration and continuity management

As mentioned, the technical and domain experts are not always available. Yet a solid understanding of the information standards' architecture and data models are required to ensure enhancements and avoid defects in the final solution as well as to preserve the continuity of the standard development. From a technical point of view, a lack of understanding in integrating the different elements of the standard results in inelegant solutions and quality problems. Quality issues can be detected through data check tools, but these are not infallible. On the contrary, problems of inelegance rely on the knowledge of the human resource conducting the work. Technical experts need to have an understanding of the toolset as well as the standard development process and industry domains.

Another cause for poor quality lies in a lack of an adequate toolchain. On one hand, most existing tools focus on software development and do not have functionalities to correctly manage or integrate data models. On the other hand, development teams rely on various tools to work on the different aspects of the solution lifecycle (e.g., planning, development, testing) and the different components of the solution such as information models or code. Having disconnected tool-chain results in a significant reliance on manual integration and process controls for project management and development tasks. Working with inadequate toolchain causes ineffective collaboration and poor version control, and manually introduced errors [6].

1.4 Conclusion

Information standards are a key enabler to digital transformation, as they provide a common language to support compatibility and interoperability between processes, assets, and stakeholders. Information standards facilitate the digitalisation of industry by increasing productivity and efficiency. Information standards promote technology adoption and interoperability, which is critical for promoting and accelerating innovation. ISO 10303, also known as STEP, is one of the oldest standards still in development and a key tool for the digitisation of manufacturing.

The standards development process is long, irregular, and difficult to plan due to a massive administrative overhead, the voluntary contribution from experts, and the irregular availability of the

organisations' resources. We have identified four main challenges, as shown in Table 2.1, which impact the implementation of standard information by companies.

Table 1.1 – Summary of the challenges of the standards development process

Viewpoints	Objectives	Causes
Requirements management	To support change management	• The prevailing use of the waterfall methodology as project management • Lack of formalisation in the definition of the requirements
Consensus management	To support consensus traceability	• Lack of detailed meetings minutes • Use of inadequate tools as collaboration/consensus tracking
Human resource commitment and management	To coordinate and optimise development planning	• Volunteer staff (recruitment, retention, and availability) • Low developer resources • Virtual distributed teams
Integration and continuity management	To improve development process quality	• Lack of knowledge • Use of inadequate toolchain

Information standards enhance the communication and interpretation of data within and across organisations. However, the adoption of standards is slowed down due to the complexity of their development process and their lifecycle. We worked on improving the standards development process through resource management and resource traceability.

CHAPTER 2: RESOURCE MANAGEMENT

This chapter focuses on resource management and the transition from predictive to adaptive project management, as the first step in our work on improving the standards development process. Section 2.1 describes the Agile methodology, as well as the reasons for standards development teams to move to this method for project management. Then, Section 2.2 presents a test case for the delivery of STEP AP242. Finally, Section 2.3 details the benefits and challenges adopting Agile can bring to the standards development process.

2.1 Agile methodology

2.1.1 Introduction to Agile

Agile has become very popular in IT management and software development projects. Agile refers to the ability and willingness to create and respond to change, as well as to be flexible. Agile is an iterative and incremental approach to project management which means the project phases are repeated multiple times throughout the project's life cycle (e.g., Figure 2.1). This approach breaks down the project scope into smaller and more manageable increments that are delivered at the end of each iteration. Each iteration takes place over a fixed-length time interval and is a repetition of the same phases (plan, design, development, test, deploy, review and launch). During each iteration, the team selects a subset of all the project's activities and does all the work required to complete the activities of this subset. At the end of each iteration, the team delivers a functional product increment and gets feedback from stakeholders. This iterative strategy allows the project to be adaptive to change and make the necessary adjustments between iterations following the stakeholders' feedback.

Agile project management is based on the Agile Manifesto, which is a collection of 4 values and 12 principles that outline the mindset that all Agile teams should aim for. These values and principles are the foundation of high-performing teams, and every team member should strive to live by them to take full advantage of the benefits of Agile.

Agile is based on short iterative release cycles, which requires more transparency and visibility, allows stakeholders to be more involved and shows them the project's progress on a regular basis [30]. Agile is mainly used to achieve high-quality projects in short periods, with better collaboration between all stakeholders, and reduced time to market [31,32]. Agile offers core benefits such as adaptability, visibility, better quality, higher customer satisfaction, early identification of problems, and reduced risks.

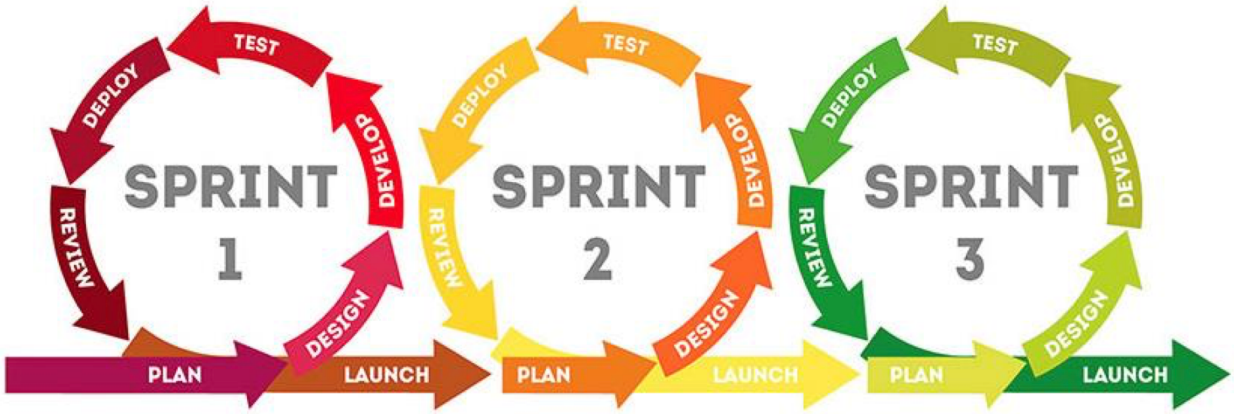

Figure 2.1 – Agile development process [33]

2.1.2 Alternative to the waterfall model

As mentioned in Section 1.3.2, standards are mainly developed using waterfall approaches for project management. An alternative to traditional project management methods (also known as predictive or waterfall approach) is the use of adaptive project management (also known as Agile) [33,34]. Predictive projects are characterised by 1) all requirements being well defined in the planning phase, 2) delivery of the final products at the end of the project, 3) absence of feedback from the stakeholders during the implementation phase, 4) leaving almost no room for unexpected changes. In contrast, Agile projects are characterised by 1) defining and changing requirements before and during the iteration, 2) frequent deliveries of functional product increments, 3) implementing the feedback provided by the stakeholders in following iterations, 4) embracing changes. Table 3.1 presents some of the key metrics that illustrate the differences between the two project management methodologies.

Table 2.1 – Comparison between the Waterfall and Agile methodologies [35]

Metrics	Waterfall methodology	Agile methodology
Development direction	Plan-driven	Change-driven
Built-in flexibility to discover problems	None	Possibility to adapt during and after each iteration
Delivery time for a usable product	Long, at the end of the project	Short, at the end of each iteration
Time-to-market and Return of Investment	Fixed, at the end of the project	Adjustable, early increments might start bringing revenue
Requirements	They are defined at the beginning of the project. These	Likely to change during the implementation phase based on

	are unlikely to change during the implementation phase	end-user feedback
Scope changes	There is a lengthy process to follow if scope changes are introduced during execution	Scope changes are welcomed
Project delivery	One delivery, at the end of the project	Incremental deliveries, at the end of each iteration
Project Risk	High at the end of the project since this is the point when the product is delivered to the market	Low at the end of the project because a shippable increment is delivered with each iteration
Collaboration between the customer and product development teams	Low	High, they can be part of the project team

The Agile and Waterfall methodologies are mainly used in software engineering and most papers comparing the methodologies use software projects as an example. According to the 2020 CHAOS Report by the Standish Group, Agile projects are more successful and present fewer challenges and failures than Waterfall projects [36]. Agile projects are three times more likely to be successful (42% for Agile against 13% for waterfall), whereas waterfall projects are two times more likely to fail (28% for waterfall against 11% for Agile) [36]. Project success rates depend on the project size. Smaller projects are more likely to succeed than larger projects. The report states that when it comes to Agile projects, small projects are three times more likely to succeed than large ones, while for waterfall projects, small projects are (almost) six times more likely to succeed than large ones [36]. However, when comparing the two types of methodologies for large projects, they estimate that the Agile approach is two times more likely to succeed.

One reason that may explain the difference in success rate between Agile and Waterfall is a high project risk due to a lack of visibility on the final product. When employing the Waterfall methodology to deliver software projects, the risk is high at the end of the project when the project's outcome is delivered [35,37]. The team concentrates on implementing the requirements defined at the beginning of the project and the product is only tested (by the team) once the implementation is done. When using the Agile methodology, the risk is lower at the end of the project because the team delivers a working product increment with each iteration [35,37]. The team receives stakeholders' feedback on the increment and can make adjustments based on them in subsequent iterations.

According to the 15th State of Agile report, the three most important reasons for adopting Agile are 1) enhancing the ability to manage changing priorities, 2) accelerating software delivery and 3) increasing

team productivity [38]. Besides, according to research by the Institute of Electrical and Electronics Engineers, “a majority of respondents' organisational units are using Agile and/or lean methods (58%)” [39]. These papers support the trend to increase the development rate. As the industry strives toward enabling the digital threads for their enterprises, rapid incremental development capabilities are required. Standard development teams still use traditional methods to create their products. These traditional methods lead to long phases of requirements documentation, product development, integration, review, and publication. Several organisations have adopted Agile to shorten the development cycle and provide a usable product to the users faster. A study analysed over 300 articles and related benefits showing that in the areas of Schedule, Productivity, Quality, and overall ROI benefits of Agile methods overshadowed that of traditional methods [40].

Since the creation in 2001 of the Agile Manifesto, there have been numerous related implementations and development of new methods. As mentioned, the Agile manifesto defines 12 principles, but three of them are particularly relevant to the development of model-based standards [39,41].

- “Deliver working software frequently, from a couple of weeks to a couple of months, with a preference to the shorter timescale” - delivering product in small, frequent increments gives stakeholders—including customers—time and regular opportunity to provide feedback on its progress. This ensures that the team does not waste too much time going in the wrong direction.
- “Working software is the primary measure of progress” - the most common technique for Agile teams to demonstrate meaningful work completion is to show a working piece of the solution. In software projects, that might be a functional piece of software. In non-software teams, this could be a critical portion of the solution that is ready to be shown to users or stakeholders in order to receive feedback.
- “At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behaviour accordingly” - in Agile, it is important to recognise that learning from successes and failures is a continual process. Teams should reflect on all of the different aspects (i.e., mistakes, challenges, trials, and triumphs) of their activities so that they can make necessary adjustments.

Note that the term “software” can be replaced by any element such as “data models” or “published data standards”.

Among the best-known Agile methods are Scrum, Kanban, Extreme Programming (XP), and Lean Development [42]. Agile methods include, but are not limited to, practices such as Backlog Management, and Continuous Delivery [43].

2.2 Test case for the delivery of STEP AP242 edition 2

Standard development generally consists of multiple teams working in tandem, in multiple locations and over a long period of time, on different aspects or capabilities of the standard, with different cadences, timeframes and stakeholders' availability. The different teams need to efficiently integrate their various deliverables together. Agile emphasises collaboration and transparency between the different teams by facilitating continuous integration and deployment. It also consists of quicker development iterations which support flexibility and make it easier to adapt to the different team members' availability.

2.2.1 Context

ISO 10303 - Application Protocol 242, also known as AP242, was published in 2014 and is the merging of two existing application protocols: Aerospace's AP203 (Configuration controlled 3D designs of mechanical parts and assemblies) and Automotive's AP214 (Core data for automotive mechanical design processes) [50]. The STEP AP242 "Managed model-based 3D engineering" defines the use of PMI (geometric dimensioning and tolerances, annotation, symbols) as a semantic representation for product definition [51,52]. It aims to define a model that would represent product and manufacturing information as well as other external references to elements or parts [51].

AP242 has a modular architecture, allowing it to evolve and expand more flexibly. The first edition of AP242 was designed to successfully merge AP203 and AP214 as well as to expand them with a set of additional capabilities such as PMI and external references on elements [50,53]. The second edition of AP242 started in September 2014 and the standard was published in 2020. This edition aimed to extend the first edition to the electrical design domain and introduce enhancements in domains such as Product Data Management (PDM), 3D PMI (e.g., Geometric Dimensions and Tolerances), composite and mechanical design [50,54].

With members of the ISO TC184/SC4 team/committee, we worked on a test case using Agile to manage the delivery cycles of AP242 edition 2, through the implementation of some elements of the SAFe framework by Scaled Agile. This work aimed to determine the benefits that implementing Agile could bring to the AP242 development and more generally to the development of standards. Furthermore, this work aligned with ISO TC184/SC4's decision to publish future editions of AP242 in shorter times and use industry requirements as a driving force [55]. Following all this, Agile has been officially adopted in 2021 for the development of the capabilities of the next edition.

2.2.2 Scaled Agile Framework (SAFe)

In addition to the most common Agile methods, there are overall frameworks that assist tie them all together to help large organisations implement at different scales. These frameworks include Scaled Agile (SAFe) [44], Disciplined Agile Delivery (DAD) [45], Large-scale Scrum (LeSS) [46], and Scrum of Scrums [47]. While some have criticised SAFe for being too prescriptive, it continues to be the most popular scalable framework by significantly outperforming Scrum of Scrums, the next nearest scaling method [38]. SAFe has experienced more industry implementations than Scrum of Scrum, LeSS, and DAD [48]. Over the last several years, there has been a growing awareness of both the opportunities and challenges offered by scaling Agile practices across the organisation.

SAFe is a Lean-Agile scaling framework that incorporates concepts from Kanban, Scrum, Extreme Programming (XP), DevOps, and Design Thinking. The first principle of SAFe is "take an economic view," which prioritises the goal of delivering value above everything else. SAFe supports collaboration of large cross-functional teams and provide a comprehensive configuration for continuous integration and deployment. The framework organises all work and teams into “Agile Release Trains” based on value streams such as sales. Although the SAFe framework is mature and gives extensive guidance on all aspects of implementing SAFe, some elements are more critical than others.

Like most Agile practices, SAFe is built on a set of core values [49]:

- **Alignment:** Synchronise the planning and execution of SAFe activities at all levels of the organisation.
- **Built-in Quality:** Build quality into all stages of solution development.
- **Transparency:** Make execution activities visible at all levels to build trust among teams and across the organisation.
- **Program Execution:** Focus on working systems and business outcomes.
- **Leadership:** Model the values and principles of SAFe.

While the SAFe framework offers many methods to implement Agile, few can bring benefit to the information standards development teams such as Backlog Management, Agile Release Trains, Program Increment Planning and Continuous Delivery Pipeline.

2.2.3 Backlog Management

A product backlog is a list of work items (e.g., user stories, bugs, changes) used by a team to coordinate work to be done in order to achieve a specific outcome [56]. Team members will develop new features, change existing functionalities, and address bugs based on the items at the top of the backlog. For

many Agile teams, the backlog is the principal project management tool. In STEP development, most teams utilise a system like Bugzilla to keep track of all the bugs.

A product backlog lists all the items required to complete the entire project. For each iteration, the team assigns items to the next sprint/milestone, and thus, it gradually reduces the number of tasks on the backlog. The team should regularly perform backlog reviews throughout the project, which will help have reliable work that is ready for the team to assign to a sprint.

There are some steps a team can take to actively optimise the management of their backlog such as defining a product owner/manager role and establishing a priority ranking. The product owner will be responsible for the why, when, and what of the product that the development team will deliver. Each team should have a person designated in this role that actively manages the backlog by reprioritising, adjusting, grooming, and adding to it. Regular prioritisation and grooming of the backlog are two main aspects of backlog management. These two activities would ensure that the team will create a product that brings value to the customers and that the backlog stays healthy by preventing it from getting too big or out of date.

2.2.4 Agile Release Trains

The Scaled Agile Framework allows teams to remain Agile even while working with larger teams and heavier workloads. SAFe provides a practice, called Agile Release Trains (ART), to bring together multiple Agile teams and align them to a common goal in order to maximise value. According to the Scaled Agile definition and framework, an Agile Release Train is used to group Agile teams that operate to develop and deliver "one or more solutions in a value stream" [57]. The ART is a virtual organisation that collaborates on planning, development, and deployment. ARTs are composed of cross-functional teams and all the capabilities such as software, hardware, firmware, and others, required to define, build, validate and release solutions. Hence, an ART is essentially a group of Teams responsible for releasing features and business values on a regular basis. All the teams of an ART are united through a common scope, program backlog, and roadmap. The ART is led by a Release Train Engineer (RTE), but it also includes other important roles such as a Product Manager, System Arch, and Business owners/Customers.

In the case of the development of Model-Based Standards like ISO 10303 Application Protocols, an ART can be used to create or revise an edition of an AP like a new edition to AP242 for example. Then, as represented in Figure 2.2, for each domain, an Agile team can be created to deliver specific capabilities, such as Electrical Wire Interconnect Systems (EWIS), Product Manufacturing Information (PMI), or Additive Manufacturing (AM). Each of these Agile teams would include a Scrum Master, a Product Owner, and a group of developers.

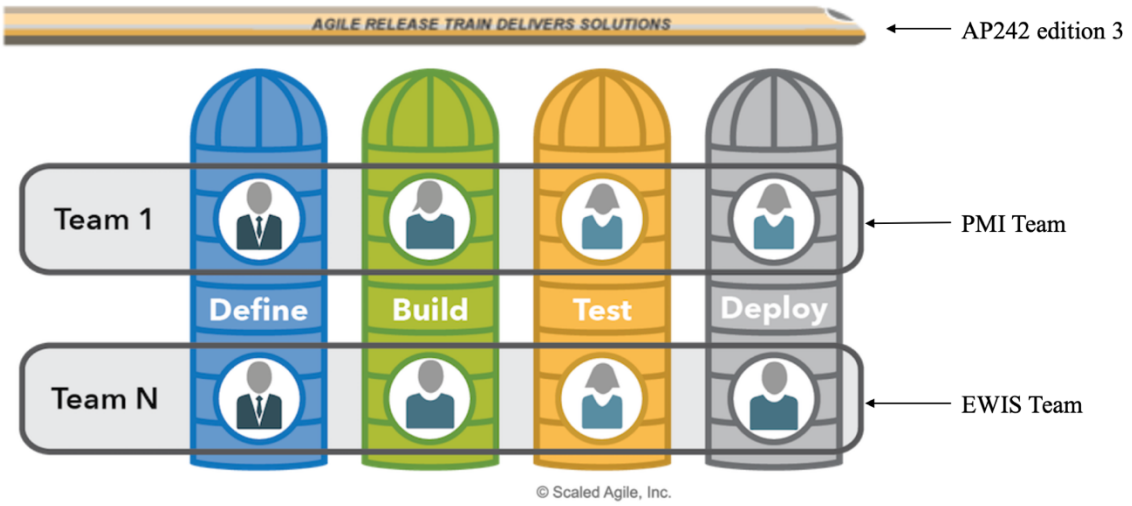

Figure 2.2 – Agile release train diagram

An ART can address one of the biggest problems with multiple teams which is the integration of their different contributions. AP242 edition 2 teams had some integration issues that could have been avoided if development iterations had been synchronised and integrated. Although each team may adopt Agile, they can and sometimes do operate with different velocities and do not sprint together. The ART solves this problem by adopting systems thinking and implementing an operating cadence and synchronisation that allows all teams to sprint together while integrating. There is no limit to how many trains can operate simultaneously. As shown in Figure 2.3, teams could use multiple trains to manage the concurrent development of AP242 edition 2, AP239 edition 3, and AP243 edition 1.

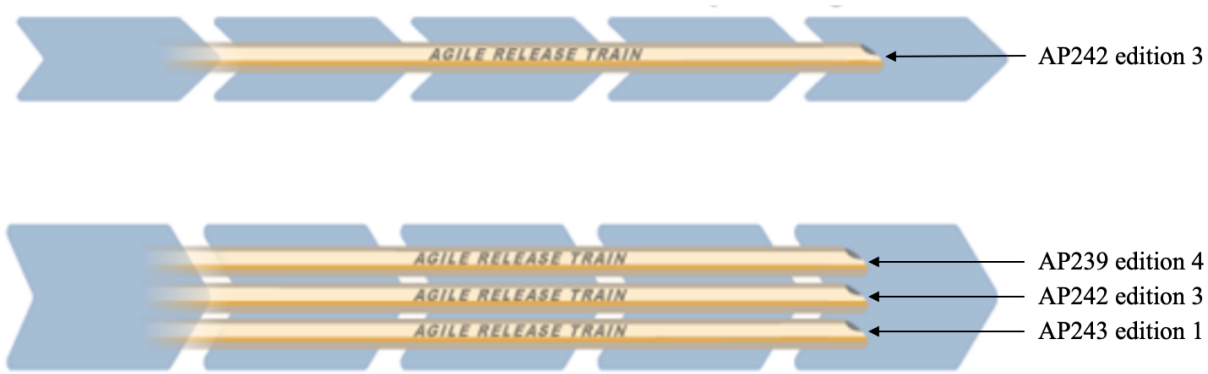

Figure 2.3 – Agile release train example

2.2.5 Program Increment Planning

In the Scaled Agile framework, a Program Increment (PI) is a fixed time period at the end of which an ART delivers value. During a PI, the teams of an ART work on the planning, building, and validation of a product increment while demonstrating value and receiving feedback. A Program Increment can be compared to a Scrum's Sprint in that they both are defined time periods with a specific goal, but a PI concerns a much larger group than a Scrum team.

Program Increment Planning focuses on determining the direction and activities of the Release Train's next PI through a large-scale planning session. PI Planning is critical to the synchronisation of the teams on the train, that is ensure that all the teams on the train are aligned and have a defined goal and plan for the next PI. This synchronisation facilitates planning and limits work in progress in order to keep teams more focused and thus, get a higher quality of work. Besides, PI Planning can be used to provide visibility, coordinate dependencies across teams, and ensure that work is prioritised in a way that prevents teams from being blocked by other teams on the train.

The RTE and team agree on the number of iterations, sometimes called sprints, that will be performed in the Increment. All Agile teams follow the same schedule and work in unison. At the beginning of each Increment, all the teams hold a planning meeting to determine their velocity and estimate and plan work packages. For voluntary teams or teams with irregularly available resources, this planning meeting is critical to determining the team resource availability and velocity. To establish its velocity, the team can use different estimation techniques. Instead of using detailed analysis and estimation, the team should choose a method where the process is quick and relative such as Planning Poker, T-Shirt Sizes, or Dot Voting for example [58].

A Plan, Do, Check, Adjust (PDCA) activity follows each iteration, allowing the team to make modifications to the plan. At the end of the Increment, a product is demonstrated to the industrial customer, which are representatives of the industrial stakeholders. This step aligns with the principles of the Agile Manifesto. For model-based standards development, where the standard is the deliverable, the product delivered at the end of an increment can be a draft schema made available for review or even an example implementation by an implementor using one of the Interoperability Forums.

2.2.6 Continuous Delivery Pipeline

One of the most important best practices of Agile, and specifically the SAFe framework, is Continuous Delivery. Continuous Delivery is largely used in software engineering and is defined as an approach in which teams produce functional software in short cycles while ensuring it can be reliably

released at any time. [59,60]. The Continuous Delivery Pipeline (CDP) is composed of three distinct yet related phases: Continuous Exploration (CE), Continuous Integration (CI), and Continuous Deployment (CD).

2.2.6.1 Continuous Exploration

Continuous Exploration is a concept that describes the processes of constantly exploring opportunities, bringing innovation, and aligning on what the product's future functionalities will be. Scaled Agile defined Continuous Exploration as a "process that fosters innovation and builds alignment on what should be built" [61]. Continuous Exploration is used by product management to define the vision of the product, the roadmap, and the next features to develop to address market and user needs. Customers and team members express new ideas which are then refined and prioritised in the backlog. In other words, while working on your current product increment, the team must consider the next stage or the next product increment.

Some ISO 10303 STEP standards (e.g., AP242, AP209, AP210) are developed using a version control system called CVS. The development community has been using, in recent years, Git, a version control system, that integrates with KANBANs and advanced communications tools like ChatOps. This can enable Agile teams to quickly explore new ideas and validate their ability to integrate while avoiding disruptions to the production system or branch line. Git is therefore a key enabler for Agile teams as the previous generation of source code management did not provide collaboration or separate and independent development areas.

2.2.6.2 Continuous Integration

Scaled Agile defines Continuous Integration (CI) as a process of developing and integrating into a continuous flow. This entails tasks such as developing, testing, integrating, and validating features from the backlog in a controlled environment to make them ready for deployment and release [62]. CI allows teams to speed up the delivery process by reducing integration times [63]. CI is frequently used alongside an Agile development workflow. The tasks compiled in the sprint backlog are distributed amongst the development team members for the delivery. Using CI enables these tasks to be developed separately and in parallel amongst the assigned developers. When one of the tasks is complete, a developer will add that new work to the CI system so it can be integrated with the rest of the project.

Software development best practices such as version control, automated testing, and build automation make CI possible. The CI process is traditionally managed by software tools such as Bitbucket/Bamboo, Jenkins, AWS CodePipeline, and Gitlab. To take advantage of the CI capabilities, the

development environment must change to a version control system based on Git technology. The new tools enable continuous exploration as well as continuous integration through a decentralised and distributed architecture, commit before merge capabilities, and integrated quality controls. One of the most powerful features of Git for CI is its branching capabilities. Branching facilitates collaboration and working on different tasks (e.g., new features, bug fixes, etc) simultaneously because it allows parallel development by providing an isolated environment for each task. For example, when working on a new feature, the team will create a new branch to encapsulate the changes and once the feature is fully developed, this branch will be merged into the main branch which consists of the official working version of the project.

In the context of standards development, CI will allow standards developers to receive immediate feedback on the success or failure of their commits by tying tools like EXPRESS Engine, JSDAI Compiles, Python scripts, or ANT Builds, as shown in Figure 2.4. Immediate feedback will allow developers to fix the issue in the current iteration rather than passing it to the end of the flow for someone else to fix (in which case the resource may have moved on and not be available). Another feature of some of the CI tools and really a requirement for standards development is the integration with other issues/tasks management systems. This integration allows everyone on the team to have a clear picture of the status of the project deliverables because feedback and issue/task lifecycle can generate new tasks. ISO 10303 Extended Architecture already uses Git capabilities but has not developed a continuous integration pipeline for quality and integration automation.

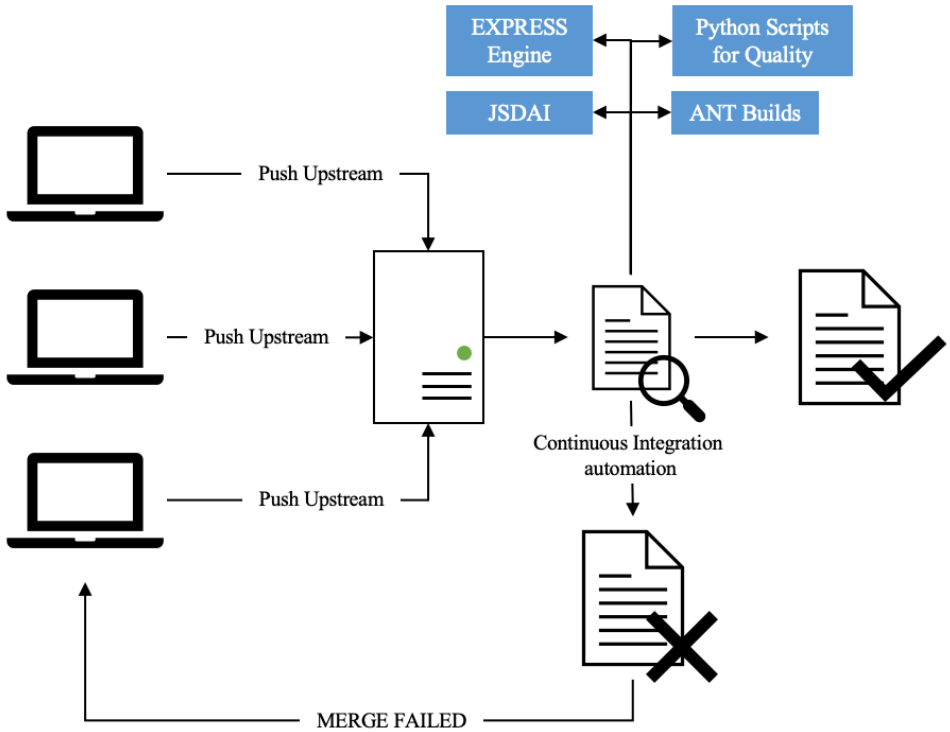

Figure 2.4 – Continuous integration

2.2.6.3 Continuous Deployment

As for Continuous Exploration and Continuous Integration, Scaled Agile offers a framework for Continuous Deployment (CD). CD is the process of deploying, verifying, and monitoring the product to ensure it is working properly and responding to issues [64]. Traditionally, the CI process is managed by software tools such as Jenkins, AWS CodeDeploy, and Bamboo.

These technologies can be used by standards development teams to automate the deployment of standards to implementer forums, formal ISO Balloting, and Publication processes and systems. If Continuous Integration is fully employed, then the deployment or publication can be performed quickly by automation. A better term for this activity in the context of a data standard could be Continuous Publication.

2.3 Conclusion on adoption of Agile for standards development

2.3.1 Benefits

Standards are mainly developed using traditional project management methods, and because of that, the development of standards takes far too long and does not align with the pace of industry innovation. Increasing the speed at which a standard is created or revised is critical to the adoption of industry standards

and the benefits that come with them. It is in this perspective that Agile offers many benefits to the standards development teams. The tools and methods offered by Agile will directly assist the developers of the standard by reducing their workflow and improving the overall quality, with the ultimate customer being the organisations in the industry implementing the capabilities of the standards sooner rather than later.

Adopting Agile for the standard development accelerates the standards publication and improves their overall quality and relevance, as they better align with the rapid changes of the industry's environment, which facilitate their adoption. Faster and wider adoption of standards make data exchange more accessible to all, which supports data traceability.

2.3.1.1 To development teams

Standards developers are one of the biggest recipients of the Agile solutions discussed previously. Agile methods and continuous delivery will bring many benefits to developers, including:

- immediate feedback loop to detect and fix issues early,
- increase transparency and visibility to other developers and team members,
- reduce tooling integration complexity,
- improve quality and testability.

For a developer, the feedback loop is probably the most part. According to a NIST research on the Impacts of Inadequate Software Testing Infrastructure, 45% of errors are discovered in the integration stage of development. A bug introduced because of poor requirements or bad coding can cost twice as much to fix in the integration stage and three times as much to fix in the testing stage. If the bug makes it to production, then it can cost up to six times as much to fix than in the stage it was introduced [65].

Teams adopting continuous delivery spend 50-70% less time on problem resolution [66]. According to another study of 34,000 open-source projects, teams that use continuous integration "release twice as often and have developers who are less worried about breaking the build" [67]. A developer will no longer be hampered by a lack of clear and complete requirements, resulting in less rework and wasted time. Because the user stories in the backlog are clearly defined and regularly updated, they can have confidence that the solution they are building meets the customer's requirements. By automating the integration and publication processes, the developers will receive immediate feedback on the quality of their work and will be able to change seamlessly. The Product Increment Planning event and backlog management will provide a clear statement of work, allowing developers to better schedule their time to support iterations.

2.3.1.2 To industry

Besides standards developers, other stakeholders (i.e., the industry) could benefit from a more Agile, consistent, and integrated standard development lifecycle. To clearly understand and identify the benefits to other stakeholders, it is important to remember their role(s) in the standard development lifecycle. Industry is both a contributor to and a user of information standards.

As a contributor, organisations satisfy the need for standards developers (technical/information modeling and domain experts), by providing funding and/or resources (i.e., experts). Reducing the complexity of the development process will alleviate the amount of time and work required by model-based standards developers. As a result, this will minimise the funding and resources required to support the design, development, publishing, and maintenance of standards. It allows industry to reduce their interoperability support costs by optimising efforts and 1) lowering their needed contribution and/or 2) speeding the delivery of standards while maintaining their level of contribution. Furthermore, an improved planning capability at the developer's level will provide contributors with a more accurate level of contribution required, facilitating their own planning.

As a user of information standards, organisations benefit from standards being widely adopted, which requires high-quality information models that are well-documented and designed to seamlessly support a wide range of implementation forms (in response to legacy, current, and future software engineering needs and trends). The benefits to the standards developers such as immediate feedback, improved quality, and testability, will lead to similar benefits in the industry. By reducing the cost of implementation and deployment in heterogeneous contexts, decluttering and streamlining the development process improves the overall quality of standards and increases adoption. Increased adoption of a standard will enable, promote, and unlock new and existing partnerships and opportunities by providing open, free, and documented formats as well as a strong, open, and neutral collaboration platform.

2.3.1.3 Summary

This test case of adopting Agile for the delivery cycle of STEP AP242 is a first step toward addressing some of the inefficiencies in the information standards lifecycle. Agile provides processes and tools to:

- Improve the management of the available resources (i.e., volunteer staff) by reducing the need for long-term planning and commitment.
- Improve collaboration between the different communities by increasing transparency and involvement of key stakeholders through a proper and faster communication channel.

- Improve the overall quality of the standard by documenting infrastructure requirements to enable automated quality control, and by allowing earlier and better feedback to detect and fix issues early in the development process.
- Finally, facilitate standard adoption by increasing the delivery rate and fastening its implementation.

2.3.2 Challenges

Despite these numerous benefits, adopting Agile for the development of standards brings some challenges that organisations need to address.

Agile is characterised by its ability to respond to change quickly. To achieve this, tasks are broken down into small increments that are delivered at the end of an iteration. Each iteration involves a team working through a full development cycle, from planning to deployment through design, implementation, and testing [31]. At the end of each iteration, the product increment developed by the team is reviewed by stakeholders and customers and changed if necessary. Adopting Agile for the standards development will entail multiple Agile iterations during one standard development cycle, which will require more effort from the development teams as well as more versioning because the standards will be implemented more often.

Requirements are crucial to the success of any project. Requirements can change throughout the project lifecycle and that is why, in Agile, the main requirements engineering activities (i.e., elicitation, analysis and negotiation, documentation, validation, and management) are repeated at each iteration [68], that is, in Agile, requirements are defined, analysed and processed multiple times (i.e., before each iteration) during the development cycle. As a result, properly capturing, tracking, and managing requirements becomes crucial, more time-consuming, and more complex.

In Agile, requirements management becomes a bigger activity. One of the main Agile artifacts is the product backlog, which is a list of the user stories that refer to the project's requirements. The content of the product backlog is regularly prioritised based on the stakeholders' feedback and changing needs. Moreover, Agile enables better visibility and transparency among the different stakeholders on the development progress, which can potentially lead to the discovery/definition of more requirements. In Agile projects, new product increments are delivered more frequently to stakeholders and customers to receive feedback and make adjustments to the next increment if necessary. Adopting Agile means that stakeholders need to get more involved and available to ensure that the team is still going in the right direction.

2.4 Conclusion

Information standards play a crucial role in the digitisation of manufacturing, and one of the main inefficiencies of the current standard development process relies on a predictive management approach.

Adopting an adaptive approach, and using Agile methodology, is an alternative to predictive management that addresses some of these inefficiencies. In this chapter, we present a test case using Agile to manage the delivery cycles of STEP AP242 edition 2 as a first step toward adopting Agile for the development of standards. Agile offers processes and tools to improve the management of the available resources, the collaboration between the different communities, and the overall quality of the standards. However, despite the benefits that Agile can bring to the standards development and adoption, it also comes with some challenges such as more involvement from the stakeholders, and better long-term management of requirements. In the next chapter, we will address some of these challenges through requirement management and resource traceability.

CHAPTER 3: RESOURCE TRACEABILITY

Transitioning from predictive to adaptive project management brings some challenges, related to requirement management and resource traceability. This chapter focuses on requirement and resource management and traceability, which is the second step in our effort for improving the standards development process. Section 3.1 describes improvements that can be made to requirements and consensus management and traceability. Section 3.2 presents existing problems in requirements engineering. Section 3.3 proposes a Requirement Elicitation model while Section 3.4 presents its implementation.

3.1 Resource management and traceability

3.1.1 Requirements management and traceability

Requirements, as well as their proper management and traceability, are critical elements to the successful execution of a project. Successful projects depend on a good understanding of the requirements and the implementation of a collaborative partnership with and between the stakeholders for requirements development and management [69]. The needs and expectations of project stakeholders must be accurately captured, documented, implemented, verified, and validated. Furthermore, ineffective requirement management is one of the primary causes of project failure and requirements issues can lead to design issues that "are more difficult and expensive to resolve" once the project development is well advanced [70].

In the standards development process, requirements come from different sources: each stakeholder has needs that the standard must address. During the development process, requirements can change according to the evolution of the stakeholders' needs and new requirements can also be created from feedback on the implemented features for example. As a result, requirements versioning and traceability should be integrated into this process to document the full lifecycle of each requirement, from its origin to its implementation. Thereby, each stakeholder can track the source of each requirement, and the changes made to these requirements and link them to the features through which they are satisfied. Tracking requirements allows stakeholders to know whether a requirement has been successfully implemented or if it needs to be reworked. Moreover, requirements management makes it easier to identify the person or group of people who issued a requirement, to get more information about it, but also offers a real-time overview of all the requirements to prioritise them.

The development of STEP began several decades ago and since that time, the stakeholders' requirements have changed in response to new business needs and the evolution of the information technologies available. In the STEP development process, ISO documents can list two types of

requirements: 1) technical requirements about the implementation form(s) of the standard, and 2) domain requirements about the environment in which the standard will be operated such as Product Manufacturing Information for example. However, it is not mandatory to provide information regarding the requirement type, issuer, or the objective behind each of them. Thus, in some cases, due to a lack of effective traceability, it is almost impossible to trace back to the concerned stakeholders to validate their requirements once the features are implemented.

Requirements traceability is a roadmap that defines where in the standard development process each requirement was implemented. Traceability can also be used to assess the impact of requirements change and expose dependencies between the requirements. On complex projects with multiple parts and different teams, like standard development, identifying these dependencies is critical yet a time-consuming and difficult process.

3.1.2 Consensus management and traceability

Meeting minutes document relevant, important, and critical topics and decisions discussed and agreed upon during online and face-to-face meetings. The archived minutes help to ensure that every member of the development process is aware of what was discussed, decided, and agreed upon. In international standard development, the physically dispersed and volunteer nature of the different teams involved makes it challenging for all the different actors to keep track of all ongoing activities and decisions made. This wide range of actors necessitates efficient tools to make it possible for all the different actors to work together. Stakeholders need transparency to be able to understand the role and the activities of everyone on the project for efficient collaboration. In this context, meeting minutes are a key communication, reporting, and traceability tool, to keep people informed and up to date with the current state of the development process. The STEP development teams host several international meetings and technical workshops during which the stakeholders meet to discuss past, current, and future developments. Standards' stakeholders cannot always attend all the international meetings and rely on meeting minutes shared with the participants, often by emails, to be and stay informed.

3.1.3 Current Tool Viability

A key contributing factor to rework in standards development and the overall extended duration of the development projects is a lack of clearly defined and traceable requirements. In the STEP community, requirements traditionally have been captured in spreadsheets and documents, and traceability is managed through inappropriate tools, such as Bugzilla, which were not designed for this type of work and lack integration with the documents that describe requirements. These methods are limited and fragmented. The growing importance of requirements management has led to the development of dedicated requirements

management and traceability tools that address these concerns and provide more comprehensive solutions, as well as the possibility to mitigate associated rework.

Several requirements management tools are available and specialise in work management including tasks, user stories, bugs, versioning, and defects tracking within a single and collaborative environment. These are all enhancements over the existing toolchain capability and consistency; nonetheless, these tools often represent requirements in plain text and do not meet the STEP community's needs for more advanced requirements management. To properly manage and validate them, the requirements must be formally defined in a semantically computer interpretable way that enables the use of reasoning tools to automate consistency checking, validation, and logical prioritisation of the requirements. A requirements management solution for information standards such as STEP must support traceability of the requirements against elements in the information models, documents, and deliverables to be able to verify that the requirements are correctly met.

Concerning meeting minutes, they are currently not consistent and easily accessible by all of the STEP community. Due to a lack of consistency and a formal template, it is impossible to link the decisions and the actions taken during the meetings to the task management software and issue log. Minutes should be detailed, formalised, and accessible to report the Five W's (i.e., who, what, when, where, and why) of each decision discussed during a meeting. A formal model should facilitate the understanding of the meeting's conduct for those who were unable to attend.

Finally, integrating Agile principles with requirements and minutes management into the standard development process requires the use of multiple tools in addition to the many different tools used to develop, implement, and maintain an information standard. Working with so many different tools and technologies means that the development team needs to ensure that there is a proper integration model in place. In the STEP environment, the tools integration situation is even more difficult due to STEP complexity, lifespan, and the use of custom tools and technologies. The community is faced with two integration challenges: 1) the requirement to migrate legacy data, and 2) the necessity to integrate standard development tools and Agile management tools, which should be able to seamlessly collaborate and exchange information.

3.2 Requirements engineering

Adopting Agile for the standards development process means that the development teams can better adapt and adjust to the changing needs of the industry. This adaptability is made possible through continuous and regular requirements changes and prioritisation. Thus, requirements management is an important activity in Agile projects.

Requirements engineering (RE) plays a crucial role in systems development [71]. Implementing requirements engineering processes contributes significantly to the success and quality of the project. Requirements engineering is an important phase in software development because it establishes solid foundations for projects and thus the project's team to successfully develop high-quality projects [72]. According to [73], the major factors of a project's success are user involvement, executive management support, and a clear statement of requirements. This report also states that the main factors causing projects to be challenged (i.e., completed projects but over-budget and over-scheduled, with fewer features than initially intended) or cancelled are lack of user involvement, incomplete requirements, and changing requirements [73]. Requirements play a critical role in project success.

3.2.1 Introduction to Requirement Engineering

Requirements engineering is defined as the process of identifying, documenting, and managing a project's requirements throughout its entire development life cycle [74,75]. Requirements engineering processes include phases that help with identifying customers' needs, defining project requirements, analysing and validating those requirements, verifying their feasibility and managing changes [76,77]. The purpose of requirements engineering is to ensure that the development of the project satisfies customers' and stakeholders' needs [78].

The requirements engineering process varies depending on several factors such as the type of the project, as well as the size and culture of the companies involved. However, there are five primary activities to all RE processes: elicitation, analysis and negotiation, documentation, validation, and management [77,79,80].

During the elicitation phase, stakeholders are consulted to understand the problem and the context of the project, as well as to collect the needs of the stakeholders and customers in order to create a list of requirements [81]. Then, the requirements are analysed to ensure their consistency, completeness, and feasibility [74]. Generally, the analysis activities result in the identification of inconsistency and conflicts between requirements [81]. Discussions and negotiations are made with the stakeholders to resolve the possible conflicts and determine a complete set of requirements for the project. Following this phase, the requirements are formally organised and written down in a document using specification mechanisms [77]. This document is used as a foundation for version control and traceability. Requirements validation is used to ensure the quality of the requirements and documentation by controlling that the requirements are complete, feasible, and correctly meet the stakeholders' needs [82]. Finally, the purpose of requirements management is to control and trace requirements progress and change [74]. This phase focuses on the

changes in requirements that reflect the changes in the stakeholders' perspectives and in the project scope [83].

3.2.2 Traditional requirements engineering challenges

Requirements engineering is a complex process, mainly because it has a direct and critical impact on the outcome of the project, but also because it requires the collaboration of a set of diverse stakeholders with varied areas of expertise and different needs [84,85]. Stakeholders have numerous and diverse needs [86]. However, projects present constraints of time and budget, which means that all of these needs cannot always be met, and as a result, it is necessary to choose the most valuable needs to transform them into requirements [87,88]. The requirements are created from needs expressed by the stakeholders and the industry, but their needs are not always clear, complete, or well-organised [86,89,90]. Unclear, incomplete, ambiguous, and conflicting requirements are identified as a common challenge of requirement engineering. The stakeholders do not always know exactly what they want and sometimes they may have unrealistic expectations that do not fit the project's budget and schedule [90]. As a result, the requirements do not all have the same level of detail, and they can be difficult to describe in a way that is understandable by all the actors involved in the project. Besides, effective communication among the different stakeholders enables a better RE process by identifying, agreeing upon, and communicating requirements more efficiently [75,91,92]. Moreover, the requirements change throughout the lifecycle of the project, which can make their management complicated in some cases [89].

Requirement management is the process of managing changes to requirements, analysing the impacts of these changes, tracing the requirements, and tracking the requirements versions and status [75,93,94]. Managing requirements is a complex activity that requires a good understanding of the project, effective management of each requirement's history, and good communication among stakeholders [95,96]. Maintaining constant communication and a clear understanding of the requirements can be a challenge, especially when working with distributed teams [95,96]. Requirement management can affect the communication between the stakeholders since it can lead to problems of division of work and responsibilities among the project's team [97]. As a result, a certain level of transparency and awareness is important and necessary in requirements management to achieve effective communication within the project [98,99,100]. Each stakeholder must know the current status of the project, the status of each requirement, and project changes, as this information can have an impact on their work and the project's success.

Requirement management is a process that continues throughout the project's lifecycle because changes in the requirements are inevitable as the needs of the stakeholders are regularly evolving

[86,101,102]. The requirements change alongside the project, and they become clearer and more detailed to reflect a better understanding of the customers' needs [93]. Requirements change management is an important activity for developing an accurate and consistent product [103]. It is used to assess the impact of the requirements change on the project from a budget, schedule, and risk perspective, as well as to validate and implement change requests [84,103,104]. Changes in requirements might affect the progress of the project and can cause delays and budget overruns [97]. The effects of requirements change on the project are difficult to predict [84]. Among the most commonly cited approaches in the literature for predicting the effort required and the consequences on the project to implement a change, change impact analysis is typically performed using traceability/dependencies analysis [105,106,107]. These approaches based on requirements traceability use trace links between requirements and system artifacts to determine the requirements impacted by the change [105,108]. Nonetheless, the generation of these links and their maintenance in these traceability-based approaches, need to be automated to save time, effectively maintain these traces over time, and properly manage the implementation of changes [105,106,109].

Requirements traceability is an important activity of requirements management whose purpose is to document the life of the requirements in both forward and backward directions through the entire system lifecycle, trace the dependencies between requirements and establish the relationship between the requirements and other development artifacts [84,102,110]. Several papers distinguish two aspects of requirements traceability: pre-traceability and post-traceability, also known as forward traceability and backward traceability [102,111,112]. Pre-traceability refers to the requirements from their origin to their formal specification, whereas post-traceability represents the trace of the requirements from their specification to their implementation [113]. These two aspects of traceability are equally important, especially with all the requirements changes, to ensure that all requirements are correctly implemented in the delivered system [114,115]. Pre-traceability helps identify organisational impacts while post-traceability helps identify technical impacts [116,117].

Traceability helps maintain consistency and assess the progress and quality of the project [115,117]. When used correctly, traceability offers many benefits. However, traceability is not always implemented in projects due to organisational, technical, or even budgetary factors [118]. Establishing and maintaining traceability relationships throughout a project is an expensive and time-consuming process [119,120]. Although requirements management tools for traceability exist, a large number of organisations do not use them because they deem them inappropriate for their needs [113,117]. These organisations will prefer to use manual techniques [121], but manually capturing and maintaining traceability links requires a significant amount of time and effort [115,120]. As in requirements management, there is no single approach for implementing requirements traceability, just as there is no consensus on a more suitable

method or a tool for systems [113,115]. Requirements traceability presents different levels of satisfaction according to the projects and the organisations and therefore, the tools used to implement traceability need to adapt to the different types of projects and their constraints [113,122]. In the domain of requirements traceability, there is a genuine need to develop cost-effective traceability solutions to facilitate the implementation of requirements traceability and make it accessible to as many projects as possible. Automating the process could help effectively capture and maintain the traceability links at a lower cost and effort [117,122,123].

3.2.3 Proposed solutions based on Agile requirements engineering

Adopting Agile can have an impact on the requirements engineering activities of the projects, and therefore some of Agile RE practices can resolve some of the traditional RE challenges. Agile methods encourage high customer involvement to increase the chances of success of the projects. Customer feedback and interaction are key parts of Agile. Direct communication is an important component of Agile requirement engineering as it promotes face-to-face communication between team members and customers over detailed and complicated documentation [124,125]. The clarification of requirements and communication gaps can be addressed through frequent face-to-face communication [126,127,128], cross-functional Agile development teams, and integrated RE processes [126,128,129]. Most Agile projects implement an iterative requirements approach, that is the RE activities are repeated each iteration [68,83]. Each iteration involves discussions between the development team and the customers to clarify, detail, refine, and validate requirements [124,130]. This involvement of the customers enables clearer requirements as well as the emergence of new requirements.

Requirement management is achieved by maintaining and updating product backlog and index cards [125,130,131] and through artifacts such as user stories and prototypes [68]. One aspect of requirement management focuses on requirement changes. As explained, change management can be challenging in traditional projects, while the dynamic nature of Agile makes it highly tolerant to changes and facilitates the addition, change, and deletion of features [124,132,133]. In Agile, frequent communication between the stakeholders and constant feedback enables early and frequent validation of requirements which simplifies the implementation of change, greatly reduces the need for major changes, and decreases the cost of addressing change [125,126]. As mentioned, requirement traceability can help estimate the impact of requirements changes on the project. There are many techniques to implement requirement traceability, but these techniques are primarily designed for traditional projects [134]. Existing traceability approaches do not take into account that requirements can be defined as user stories or tests as is the case in Agile [135]. In Agile, acceptance tests and the user Stories (i.e., requirements) they verify should reference each other, because if the requirements change, then the acceptance tests associated should change too [134,136].

Moreover, several methods are developed to capture traceability links between Agile artifacts (e.g., user story, test, code) [135,137,138,139].

3.2.4 Conclusion on requirements engineering

Requirement engineering contributes to the quality and success of projects, but it also presents some challenges, some of which have been detailed in a previous section. The adoption of Agile methods can change the execution of RE activities. Agile RE brings some benefits such as better communication and involvement of the stakeholders, a clearer understanding of the requirements, and improved responsiveness to change [125,140]. It also can alleviate some of the traditional RE challenges, as shown in Table 4.1.

Table 3.1 – Summary of traditional RE challenges resolved by Agile RE practices

RE challenges	Agile RE practices
Unclear and incomplete requirements [86,89,90]	Face-to-face communication, cross-functional teams, integrated RE process [126,127,128,129]
Communication issues [91,97,100]	
Requirement change [84,97]	Responsive to change [124,132,133] Frequent communication and feedback [125,126]
Requirement traceability [113,122]	Methods adapted to Agile [135,137,138,139]

However, Agile RE can introduce new challenges to projects such as i) minimal documentation (i.e., user stories and product backlogs are the only documents), ii) negligence of non-functional requirements (i.e., users stories only describe the system and product features), iii) difficulties for requirements prioritisation (i.e., prioritisation only based on immediate business value) and iv) inadequacy of the user stories format (i.e., lack of power of expression don't convey enough information) [83,126,128,140,141].

3.3 Requirement Elicitation Model

3.3.1 Requirement Elicitation

3.3.1.1 Focus

The different phases of requirements engineering present many challenges [142], yet most of the literature focuses on conflict management [87,143,144,145,146] and predictive environment, leaving out

requirements elicitation, a key step to their management [147]. Requirements elicitation is the first phase of the requirements engineering process [148,149]. The objectives of this phase are to discover the stakeholders' needs and gather information about the project and its scope in order to understand and specify as clearly as possible the requirements of the project [84,149,150]. Requirements elicitation is a critical step in project development and is sometimes considered a complex process [76,151], as errors made at this phase can have significant consequences on the development of the project such as cost overruns, delays, poor project quality or even non-satisfaction of the end-users [76,152,153]. The elicitation phase is critical for gathering information on customer needs that could be used to trace requirements to their source and validate them later in the development process. There is no standard definition of what information should be collected for the requirement elicitation. To facilitate this process, many methods have been developed to capture and represent requirements. The solutions presented in the following section have been developed by SDOs to capture requirements.

3.3.1.2 Existing solutions

The Requirements Interchange Format, also known as ReqIF, is a standardised XML-based format for requirements exchange, driven by the automotive industry [154,155]. ReqIF is a tool-independent and non-proprietary format aimed to "represent all important features of requirements data, including requirements in the form of attributed data elements, links (traces), views on the requirements data, and permission data" [156]. It also supports the exchange of requirements, as well as their associated metadata, between requirements management tools [154]. This format meets industry needs for exchanging requirements data between different organisations without requiring them to use the same management tools [157]. ReqIF is widely used and supported by practically all existing requirements management tools. A ReqIF file is composed of a header, tool-specific extensions, and the actual requirements data that includes both the description of the data types and the data itself. The requirements data contains multiple data elements [155,156] such as i) SpecObjects that represent the requirements through a number of attributes, ii) Specifications are hierarchical structures that organise SpecObjects, iii) SpecRelations represent the relationships between SpecObjects, and iv) SpecType describes the requirement type (e.g., functional, quality, etc) as well as which attributes these requirement types can have.

The System Modeling Language (SysML) is a modeling language for the specification, analysis, design, and verification of a broad range of complex systems [158]. It enables to effectively capture information in order to facilitate integration and reuse in a larger context [159]. SysML includes three different types of diagrams: behaviour diagrams, structure diagrams, and requirement diagrams [159,160,161]. SysML diagrams are used to represent text-based requirements for a system or subsystem [162]. These diagrams can effectively capture functional as well as non-functional requirements. The

SysML requirement diagram specifies requirements in a way that bridges the gap between natural language, which can bring ambiguities and formal or semi-formal models [161,159]. This SysML diagram captures the relationships between requirements, system design models, and test cases [160]. SysML requirements are primarily defined by a unique identifier and a text-based definition, but additional properties such as priority, status, or verification status can be included [163].

3.3.1.3 Presentation of the model requirements

These models focus primarily on a limited, simplified, and textual definition of the requirements, and the relationships between them. These models were designed to support the representation and exchange of requirements, and they were not meant to support advanced requirements management and planning activities [144] in an adaptive environment. As explained previously, when moving from a waterfall to an adaptive project management approach, requirements management takes a different and more important role and requires more additional information and visibility to be properly executed [164]. Furthermore, working with distributed teams and volunteer human resources requires better traceability and visibility of both decisions and contributions. An optimal solution should focus on defining the information to capture in order to overcome some of the challenges presented in Section 1.4 and meet the following needs:

- (N1) it should ensure the traceability (and visibility) of resources by 1) associating people with work items and tasks to which they contribute or supervise (i.e., their level of engagement), 2) associating people with the meetings they attend, and 3) associating work items with meetings during which they are discussed in order to follow the progress of the work item. This would help with the resource commitment and management challenges;
- (N2) it should ensure the traceability (and visibility) of decisions by 1) associating people with the decisions they make, 2) associating decisions with the meetings during which they are made, and 3) tracking the evolution of decisions related to work items and requirements as the project progress; This would help with the issues related to consensus management;
- (N3) it should ensure requirements definition management by associating requirements with their source and a context in order to be able to validate them with their owner once completed. This would help with integration and continuity management.

3.3.2 Presentation of the model

None of the models presented in the previous section captures enough information to overcome the challenges and meet the needs we presented because they lack a (precise) definition of 1) the requirements' context and source, 2) the requirement ownership information, 3) the associated resource(s) management,

4) the associated work breakdown, 5) key decisions, and 6) tracking information between requirements and associated deliverables. As a result, we developed a requirement elicitation model [165]. This model focuses on requirement elicitation, and consensus and resources traceability. This model can be utilised on its own or as an extension to existing solutions such as the ones introduced previously (i.e., ReqIF, SysML). As an extension, this model can help extend and enrich traditional solutions that were not initially designed for adaptive management, as well as for requirements and decision traceability. Figure 3.1 illustrates this requirement elicitation model. Our model is a UML Class diagram, composed of six concepts (Meeting, Requirement, Work Item, Task, Member, and Decision) in response to our three main needs: requirements definition management and traceability, decisions traceability, and commitment and resources traceability. This model supports our needs as demonstrated in Table 4.2.

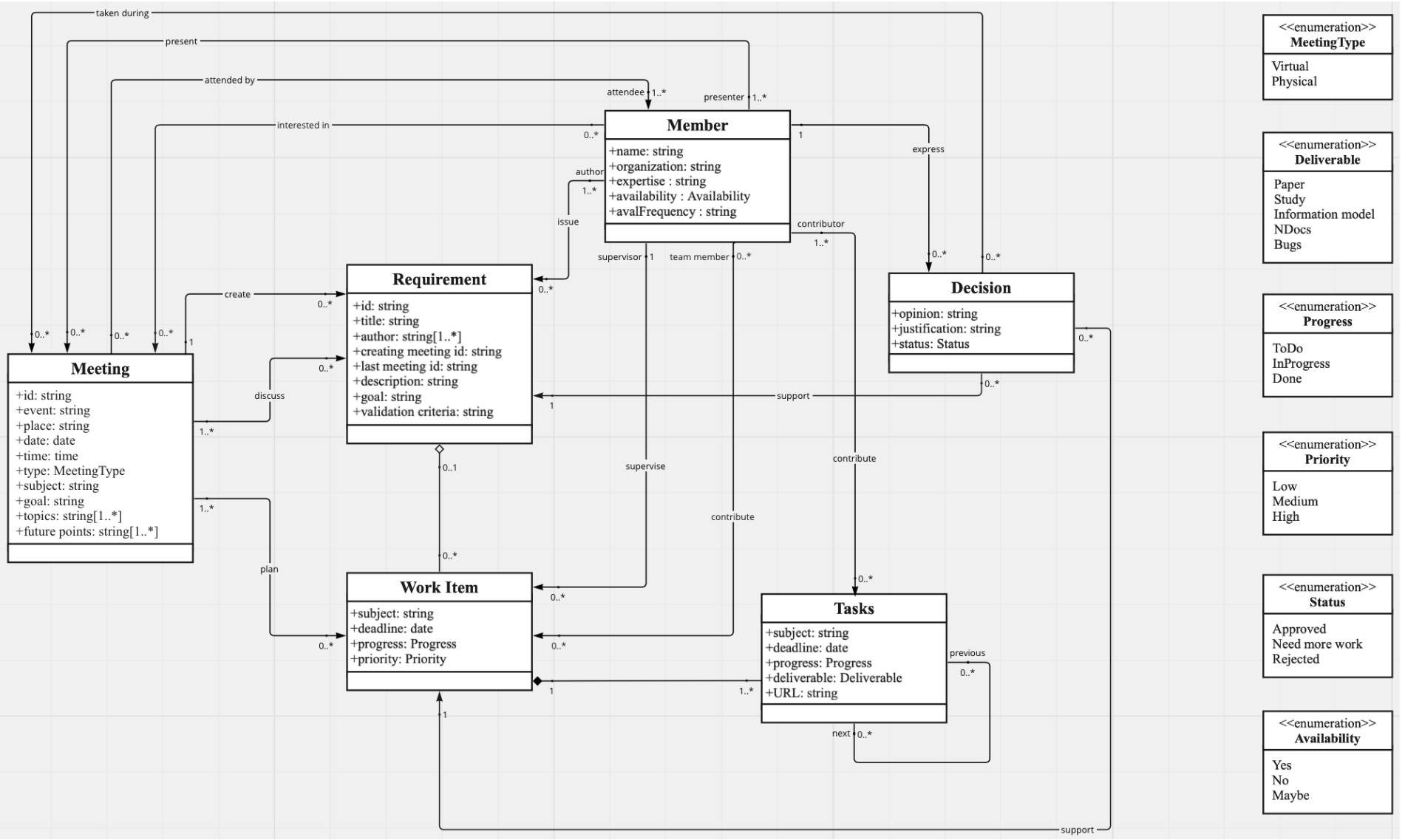

Figure 3.1 – Requirement elicitation model

Table 3.2 – List of the needs we meet and Agile RE challenges we address with the concepts from our elicitation model

Needs	Agile RE challenges (Section 3.2.4)	Model concepts
Commitment and resources traceability (N1)		Requirement, Work Item, Task, Meeting and Member concepts
Decisions traceability (N2)	Minimal documentation, inadequacy of the user stories format	Decision, Member and Meeting concepts
Requirements definition management and traceability (N3)	Minimal documentation, inadequacy of the user stories format and *difficulties for requirements prioritisation*	Member, Meeting and Requirements concepts

The objective of the proposed model is to keep track of decisions and requirements throughout the development of a project, which is crucial for long projects involving different distributed stakeholders such as standards development projects. This model can be used to determine the level of engagement of the different stakeholders and thus, better manage the available resources. Furthermore, this model helps record all decisions taken during meetings and make that information available to those who were unable to attend. It also helps ensure the traceability of requirements from their creation to their implementation in order to validate them directly with the people who had expressed them.

This model aims to provide requirements and consensus traceability in order to keep track of all information about a requirement through its lifecycle and all decisions and opinions discussed during meetings. This model can help the project teams for the prioritisation or validation of requirements and conflict management by providing information about the requirements and the decisions taken, but requirement validation, requirement prioritisation and conflict management are not in the current scope of our model.

3.3.2.1 Meeting concept

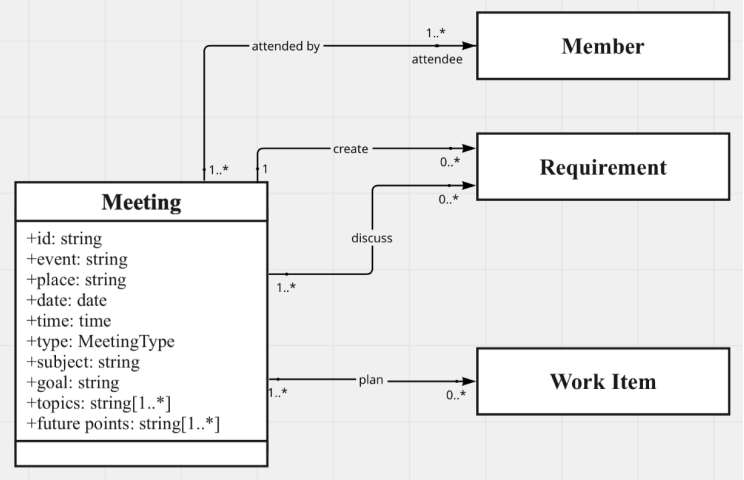

Figure 3.2 – Requirement elicitation model – Meeting concept

The Meeting concept characterises the meetings and their various properties (e.g., Figure 3.2), and it includes several properties for keeping track of all the details of the meeting.

- Each meeting is assigned a unique identifier.
- The *event* attribute specifies the event during which the meeting took place, this attribute is not mandatory because meetings are not necessarily included in a particular event. An example of an event can be an Implementer Forum or Technical Committee (e.g., ISO/TC 184/SC 4) plenary meeting.
- The location, date, and time of the meeting are respectively defined through the *place*, *date*, and *time* attributes;
- The *type* attribute indicates whether the meeting was in person or virtual. If the meeting was virtual, the *place* attribute is left unfilled;
- The *subject* attribute specifies the main subject of the meeting from a high-level point of view. An example of a subject can be technical discussion on a specific tool or WG update meeting.
- The agenda of the meeting, which is the list of the items that have been discussed during the meeting, is detailed through the *topic* attribute. An example of an agenda can be 1) recently added capabilities, 2) demonstration, or 1) status review 10303-15, 2) status review 10303-16, 3) discussions.
- The *goal* attribute represents the objective of the meeting. An example of a meeting goal can be to update users on tool development or opportunities to discuss the WG progress.
- The list of subjects that will be covered at the next meeting is specified through the *future points* attribute.

Note: all examples are in the context of STEP development.

The Meeting concept has relationships with three other concepts. Firstly, the Meeting concept is linked to the Member concept, which, in that case, represents the multiple people who attend the meetings. Secondly, the Meeting and Requirement concepts share two relationships. During a meeting, requirements can be defined, and they can also be discussed, reviewed, updated, and changed. The definition of a requirement took place during one meeting, but requirements can be discussed or changed during more than one meeting. Thirdly, the Meeting concept is linked to the Work item concept, as work items can be planned and discussed during one or more meetings.

3.3.2.2 Member concept

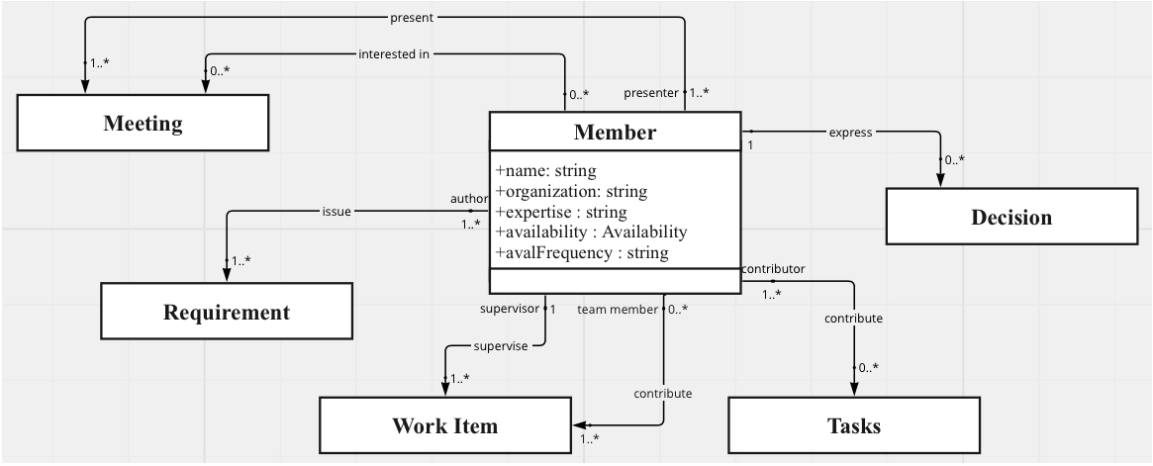

Figure 3.3 – Requirement elicitation model –Member concept

The Member concept represents the characteristics of a stakeholder (e.g., Figure 3.3). A member can be anyone involved in the project such as a manager, a developer, or a customer. A member is characterised by their name, the organisation to which that person belongs and their field of expertise. This concept also represents the availability of the member for the current iteration which aims to facilitate the scheduling and allocation of resources. A member availability is characterised by two attributes: the availability attribute that represents if the member is available, maybe or not; the avalFrequency attribute represents the frequency at which the member is available (e.g., 2h per week, 1 day per week).

The avalFrequency attribute is left unfilled if the member is unavailable. This availability applies to the agile development iteration and not the standard development cycle, because it would be too complicated for the team members to know their availability over such a period of time. Standard

development cycles last over several months, or even years, while development iterations usually last over a few weeks, depending on the teams.

The Member concept is linked to all the other concepts:

- to the Meeting concept: during a meeting, different persons can present talks as multiple topics are discussed, knowing the presenter(s) of the meeting facilitate following up for more information. The presenter(s) must be a participant in the meeting; besides, some people can be unable to attend the meeting but still be interested in what was discussed in the meeting, which is why, a member can specify its interest in having access to the meeting information for follow-up.
- to the Decision concept: any member can express an opinion on requirements or work items;
- to the Requirement concept: a requirement can have one or more authors;
- to the Work item concept, through two relationships: a work item is supervised by one person and multiple people can work on it. We considered that a supervisor also contributes to the work item that he supervises and therefore, multiple relation links between the same member and work item are unnecessary. A supervisor can oversee several work items. In the case of the supervisor is the only contributor to the work item, then there is only one link between the two concepts;
- to the Task concept: one or more people can contribute to some tasks.

3.3.2.3 Decision concept

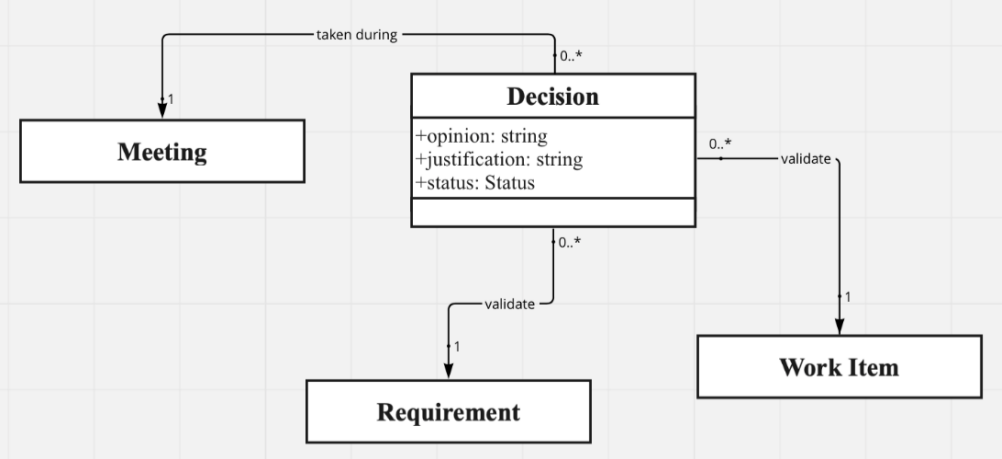

Figure 3.4 – Requirement elicitation model – Decision concept

During a meeting, each participant is free to express an opinion on the different topics discussed. The objective of the Decision concept is to keep track of the decisions made during a meeting (e.g., Figure 3.4). This concept includes three attributes to characterise these decisions. The *opinion* and *justification* attributes provide a participant's opinion on a particular requirement or work item, as well as its justification. The

last attribute represents the status of the decision and can presently only take one of three following values: Approved, Need more word or Rejected.

The Decision concept has relationships with three other concepts: the Meeting concept, the Requirement concept, and the Work Item concept. A decision is associated with the meeting in which this decision was made, as well as with the requirement or work item to which this decision relates. Several decisions can be taken during the same meeting and several decisions can be associated with one requirement or work item.

A decision represents the discussion that happened during a meeting. It is used to support a requirement or a work item by representing the opinions of the different meeting's attendees on requirements and work items. A decision does not validate the creation of a work item or a requirement or its completion, but it can be helpful for the validation process.

3.3.2.4 Requirement concept

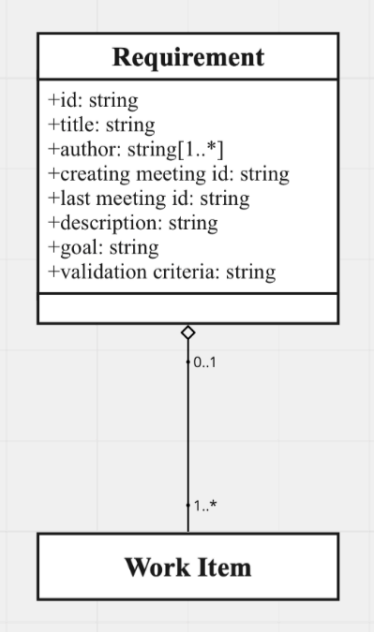

Figure 3.5 – Requirement elicitation model – Requirement concept

The Requirement concept describes the different properties that characterise a requirement (e.g., Figure 3.5). A requirement is composed of eight attributes. Firstly, a requirement is identified by a unique id. A requirement is also characterised by a title and a primary author. The Requirement concept also includes the id of the meeting during which the requirement was defined as well as the id of the last meeting during which the requirement was updated. New requirements are not always discussed during official meetings, as stakeholders may sometimes introduce requirements during informal discussions. Therefore,

we considered that the *creation meeting id* attribute can refer to the first meeting a requirement is officially brought out to the project team. When a requirement is created, the *creation meeting id* and *last meeting id* attributes are identical. Finally, a requirement is also characterised by a description, a goal, and validation criteria.

3.3.2.5 Work Item concept

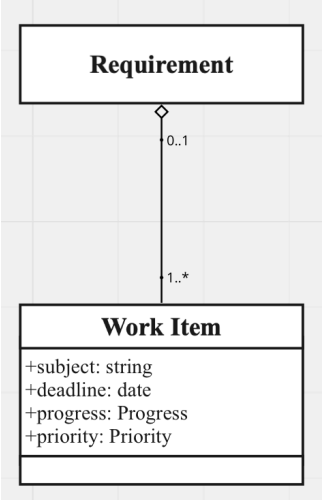

Figure 3.6 – Requirement elicitation model – Work Item concept

A project is made up of multiple work items that represent the different steps to reach the goal of the project. Work items help to plan and manage projects. In our scenario, a requirement can be coupled with one or more work items. Conversely, a work item does not necessarily depend on a requirement. The aggregation link between requirement and work item means that requirements are addressed through work items but that not every work item is linked to a requirement. A work item is characterised by a high-level subject, a deadline, a progress status, and a priority status (e.g., Figure 3.6). The progress status of a work item can be To Do, In Progress, or Done, while its priority can be Low, Medium, or High. The progress of a work item depends on the progress of the tasks that compose it. A work item cannot be marked as Done until all of its tasks are also marked as Done.

3.3.2.6 Task concept

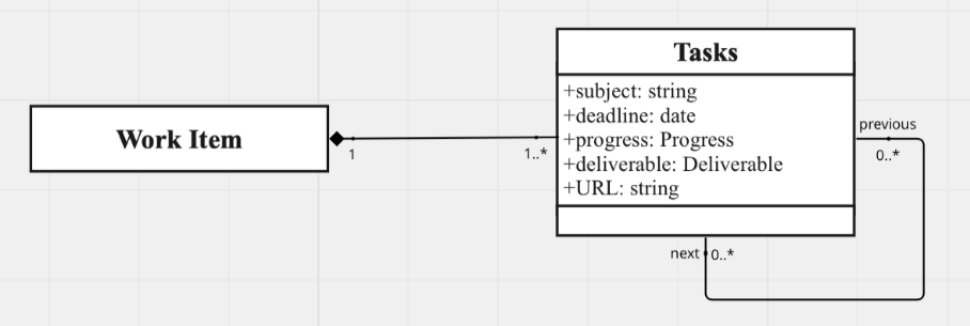

Figure 3.7 – Requirement elicitation model – Task concept

A work item is composed of one or more tasks. The Task concept can represent user stories, tasks, bugs, or issues. A task is characterised by five attributes (e.g., Figure 3.7): i) a subject that specifies the objective of the task, ii) a deadline, iii) a progress status, iv) the type of deliverable returned, and v) the URL associated with the commit corresponding to the task's deliverable. In the same way, as for the work item, the value of the task's progress includes To Do, In Progress, or Done. Currently, the deliverable of a task can be a Paper, an Information model, some NDocs, or a Bug. The *deliverable* attribute can be left empty until the task has started, giving time for the team to decide what type of deliverable they are aiming for. Concerning the *URL* attribute, software development teams are largely using Git and its commit capabilities for version control and collaborative work, and standards development teams are starting too therefore each step towards the completion of a task is associated with a commit. Yet, when the task's progress is still marked as To Do, which means that the task has not started, then the *URL* attribute is left unfilled. In some cases, the different tasks included in a work item have dependencies and must follow a specific order, which is represented by the relationship between the tasks.

3.3.2.7 Conclusion on the model

To summarise, this requirement elicitation model helps keep track of and communicate information about project meetings, decisions, requirements, and work items. This model also keeps all stakeholders up to date on the progress of the project and the many subjects discussed during the various project meetings. Moreover, this model facilitates the traceability of requirements by recording their creation and authors, as well as any changes that may have occurred and the status of the associated work items. This information is essential to better understand and follow the evolution of the requirements over time, as well as to validate them directly with the people who expressed them in the first place.

3.3.3 Example of use of the model

Figure 3.8 depicts a virtual technical meeting attended by five stakeholders. The major goal of this project is to discuss project management methods in light of the Agile transition. During this meeting, the req20_09 requirement about adopting an adaptive method for project management was revised, and two work items were discussed. The first work item, which is associated with the req20_09 requirement, focuses on the transition to Agile and is composed of four tasks. The first task is complete, the second one is still in progress but has sufficiently progressed for the third task to begin, and finally, the last task has not started yet, waiting for the previous tasks to be completed. This task does not have an associated repository or a defined deliverable. The second work item that concerns the review of the issues log is still in progress.

During this meeting, the task about the bugs classification was reviewed and completed, and their resolution was planned.

Due to its design, our solution can be used on its own or integrated into an existing model or tool. For instance, to integrate this solution to ReqIF, our Requirement concept would replace the SpecObject element of the ReqIF model, and our Work Item concept would replace the System Component element of ReqIF. As an extension, it can extend and enrich traditional solutions that were not initially designed for adaptive management.

In the next section, we suggest a possible implementation of the requirement elicitation model through the customisation of an existing work management tool.

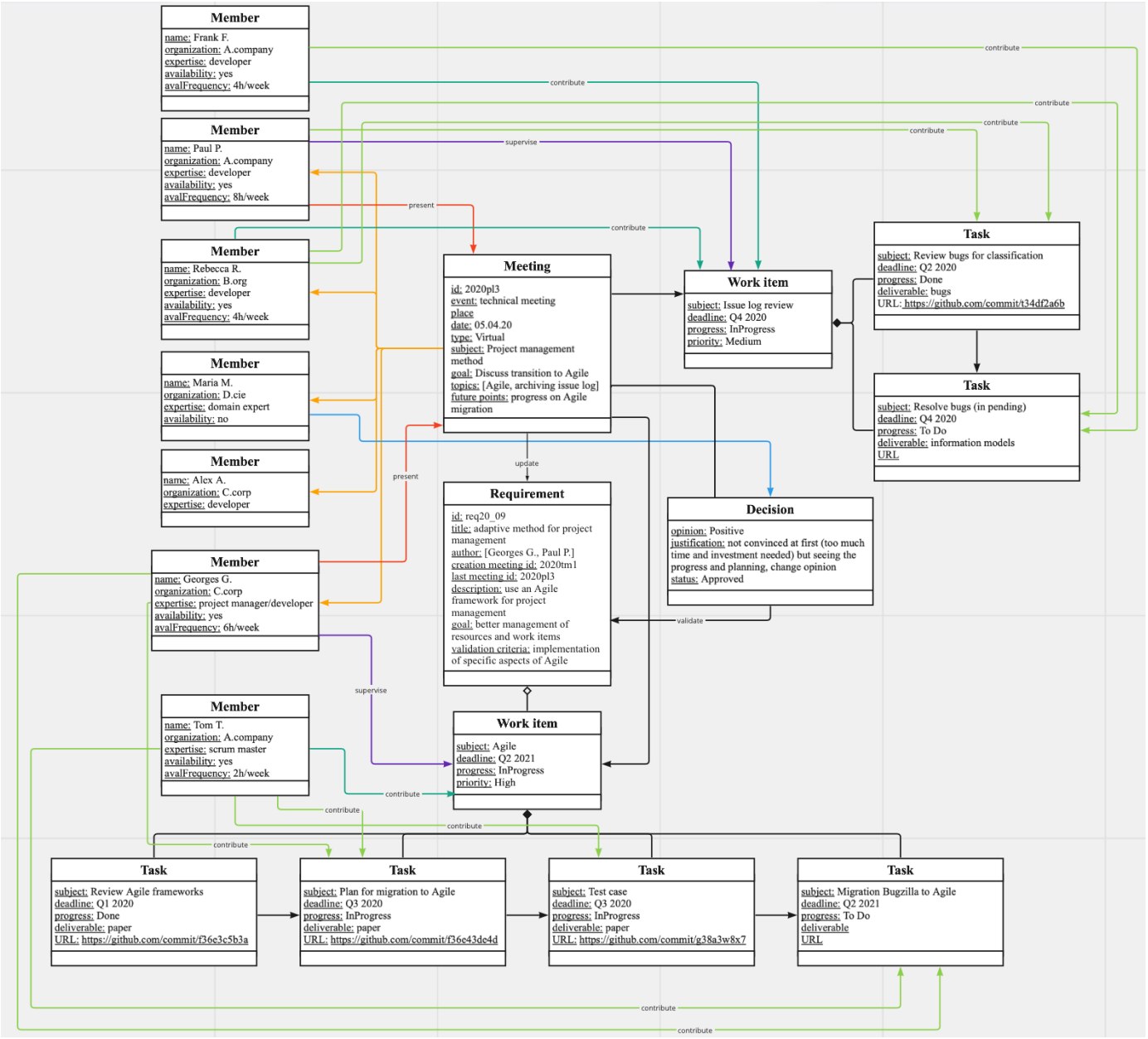

Figure 3.8 – Basic example of our Requirement Elicitation model (UML Object Diagram)

3.4 Implementation of the model

Strong market competition and demand for new innovative products lead to a growing need to build products faster yet reliably. To support product development and project management, numerous tools are created to assist with project coordination and collaboration. Among these tools, Jira is one of the most popular Agile project management tools as it offers an extensive set of features and the ability to extend these capabilities. In addition, the ISO/TC184/SC4 committee, which is responsible for the development of STEP, has officially transitioned from Bugzilla to Jira at the end of 2021 to manage issues/bugs for ISO 10303 STEP development.

Through various features, Jira provides a comprehensive suite of issue and project tracking throughout the entire product development lifecycle. Jira offers Agile project management capabilities by supporting Agile methodologies like Scrum and Kanban. Jira enables teams to 1) create backlogs to store tasks and prioritise work, 2) plan sprints, 3) define epics to organise tasks, and 4) define user stories to complement the tasks. Jira also includes roadmaps to get a high-level overview of the project, as well as reporting tools such as burn-down charts, sprint reports, and velocity charts. Furthermore, Jira provides a rich toolkit for developers. It integrates with Github and BitBucket, among other development tools, enabling traceability for all backlog items. It also includes the Jira Query Language (JQL) that allows teams to create filters and search for issues, as well as a Confluence workspace to support collaboration and share knowledge. Moreover, Jira can be customised to meet the needs of project teams for issue tracking or task management. Jira can also manage large teams with large number of requirements and issues, as it is the case in standard development.

We suggest an implementation of the requirement elicitation model through the customisation of Jira. The Requirement, Work Item, and Task concepts of the model are represented through new issue types and custom fields. The attributes of each of these concepts are mapped to fields in Jira. By default, Jira issues contain two description fields (Summary and Description) and four context fields (Status, Assignee, Labels, and Reporter), as well as various actions such as attach files, link commit, or confluence pages, add checklist or comments.

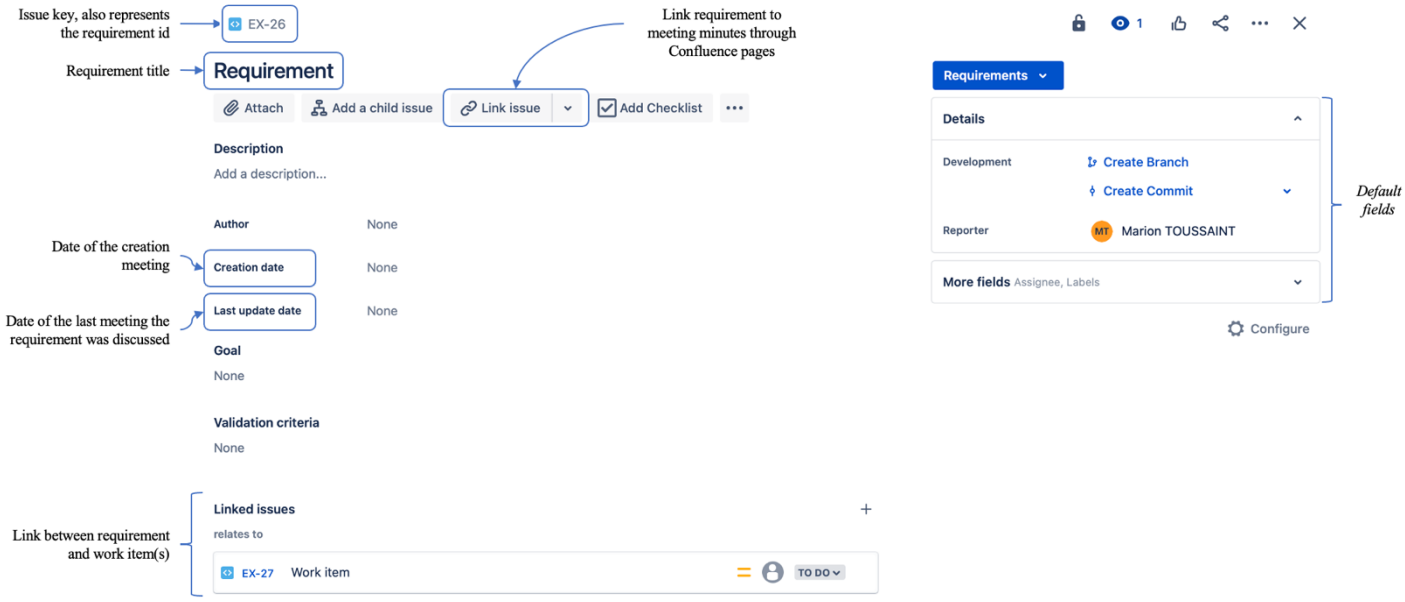

Figure 3.9 – Screenshot of the implementation of the Requirement concept (Jira)

Figure 3.9 depicts the mapping of the attributes of the Requirement concept to the fields of Requirements issue type in Jira. As shown in Figure 3.9, Jira fields were defined to represent the *id*, *name*, *author*, *goal*, and *validation criteria* attributes of the Requirement concept. Regarding the *creation meeting id* and *last meeting id* attributes, some changes have been made. First, the ids of the creation meeting and last update meeting are replaced by the dates of creation and last update meetings respectively. Second, through the *link issue*, the requirement can be linked to Confluence pages that contain the meeting minutes during which the requirement is discussed (creation or update). As mentioned, Jira integrates Confluence, which allows teams, among others, to record the meeting minutes and reference Jira issues (such as Requirement, Work Item, or Task issue types we defined). Hence, the information contained in the Meeting and Decision concepts of the model, are documented on Confluence pages. Besides, the relation between requirement and its associated work item is defined through the “Relates” link issue type.

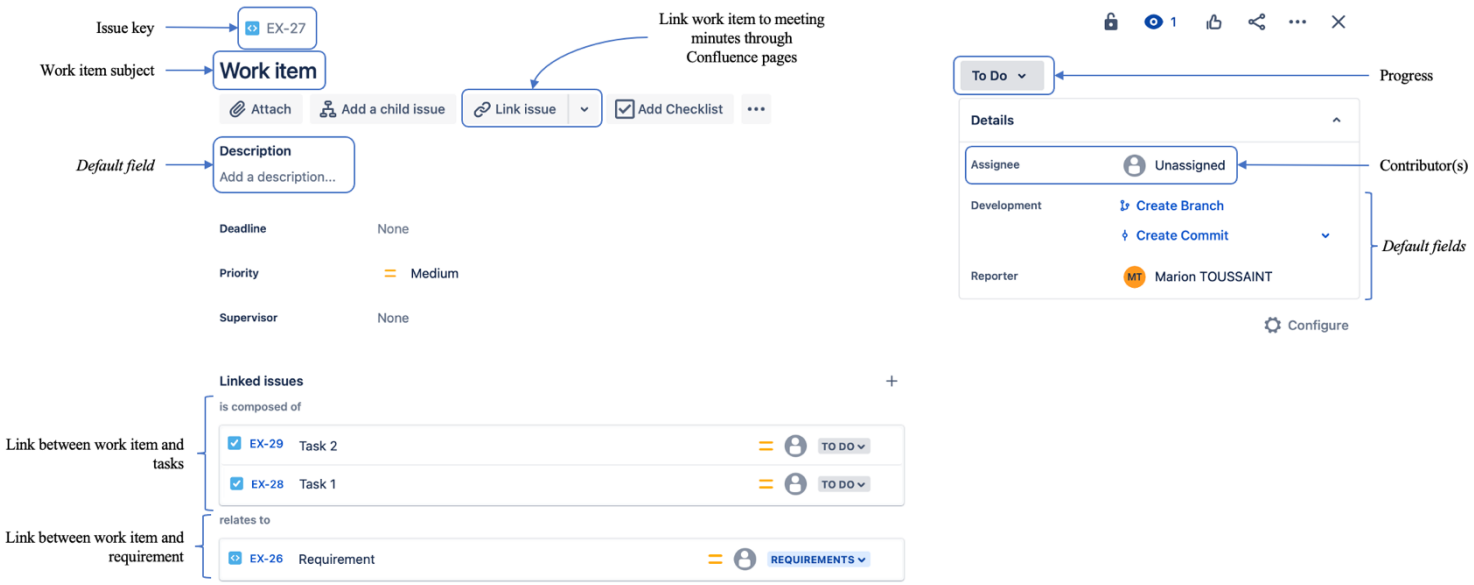

Figure 3.10 – Screenshot of the implementation of the Work Item concept (Jira)

Figure 3.10 depicts the mapping of the attributes of the Work Item concept to the fields of Work item issue type in Jira. Custom fields were defined to represent the deadline and priority attributes as well as the supervisor of the work item. Unlike the model in which the priority attribute can take one of the three values defined in the Priority enumeration, the Priority field in Jira is a dropdown list containing five values (Highest, High, Medium, Low, Lowest). The subject and progress attributes, as well as the contributors, are represented by Summary, Status, and Assignee default fields respectively. Similarly to requirements, the Confluence pages in which work items are planned or discussed can be linked to work item issues. In the Requirement elicitation model, the Work Item is linked to both Requirement and Task concepts. In the Jira implementation of the model, we use the “Relates” link issue type to represent the relation between requirement and work item, while we define the “Compose” link issue type to illustrate the link between work item and task(s).

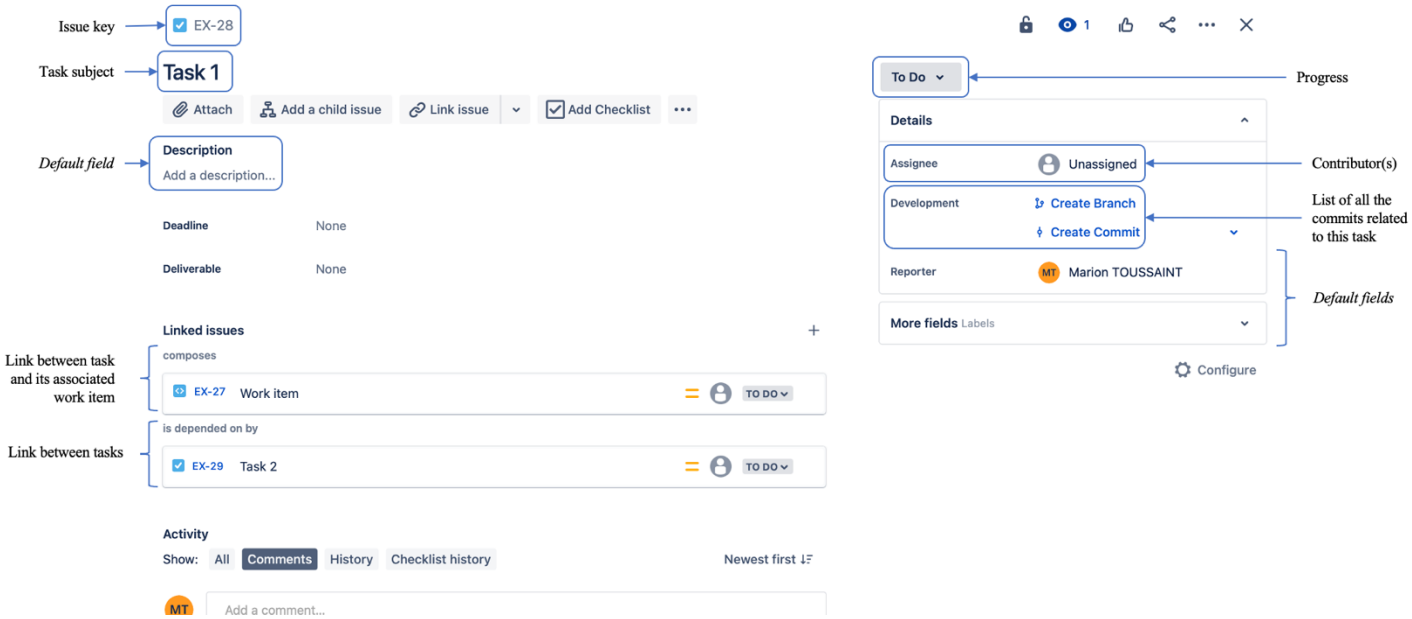

Figure 3.11 – Screenshot of the implementation of the Task concept (Jira)

Figure 3.11 depicts the mapping of the attributes of the Task concept to the fields of Task issue type in Jira. Custom fields were defined to represent the *deadline* and *deliverable* attributes, while similar to the Work item issue, the *subject* and *progress* attributes, and the contributors are represented by default fields. The relation between the work item and the task is defined through the "Compose" link issue type, as explained above. Tasks can be linked together to illustrate the order in which they must be done, which is useful if their realisation depends on the other task(s). We defined the "Depend" link issue type to represent the relationship between tasks. As mentioned previously, Jira can integrate with Github and BitBucket, and commits can be linked to Jira issues by including the issue keys in the commit message. Thus, for each issue, it is possible to see the list of commits associated with the issue as well as the date of the commit, its author, and the files related to the commit, as depicted in Figure 3.12. This functionality enables us to represent the URL attribute of the model's Task concept.

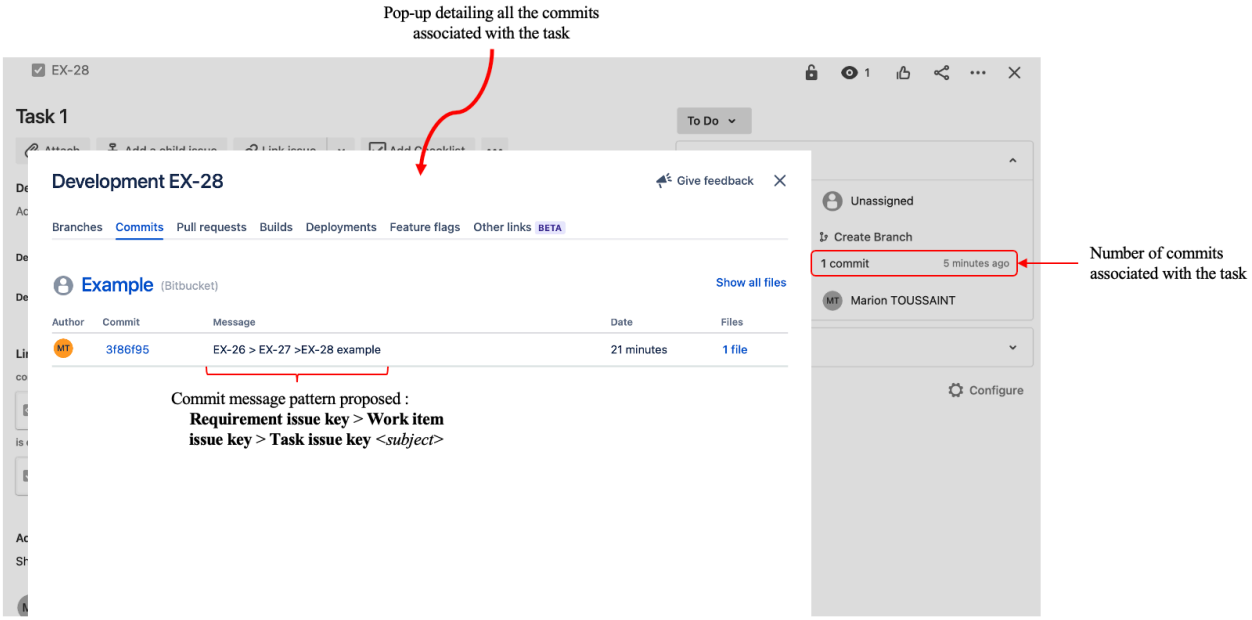

Figure 3.12 – Example of the commit message using issue keys (Jira)

In order to create full traceability between the requirements, work items, tasks, and deliverables, we propose a pattern to include in the commit message, which is **Requirement issue key** > **Work item issue key** > **Task issue key** <*subject*>. According to Figures 3.9, 3.10 and 3.11, the requirement EX-26 is related to the work item EX-27 which is composed of tasks EX-28 and EX-29. The message of a commit associated with task EX-28 would be: **EX-26** > **EX-27** > **EX-28** *subject of the commit,* as shown in Figure 3.12. In the case of work items not associated with the requirement, the pattern of the commit message is **Work item issue key** > **Task issue key** <*subject*>. Including the issue key of the requirement associated with the work items and tasks allows teams to find all the commits related to the implementation of this requirement for example. As the issue key of the requirement and work item associated with the task are included in the commit message, the commit is linked to both requirement and work item, and it is possible for each of them to access the list of all commits associated with them, in the same way as for the Task issue type (see Figure 3.12).

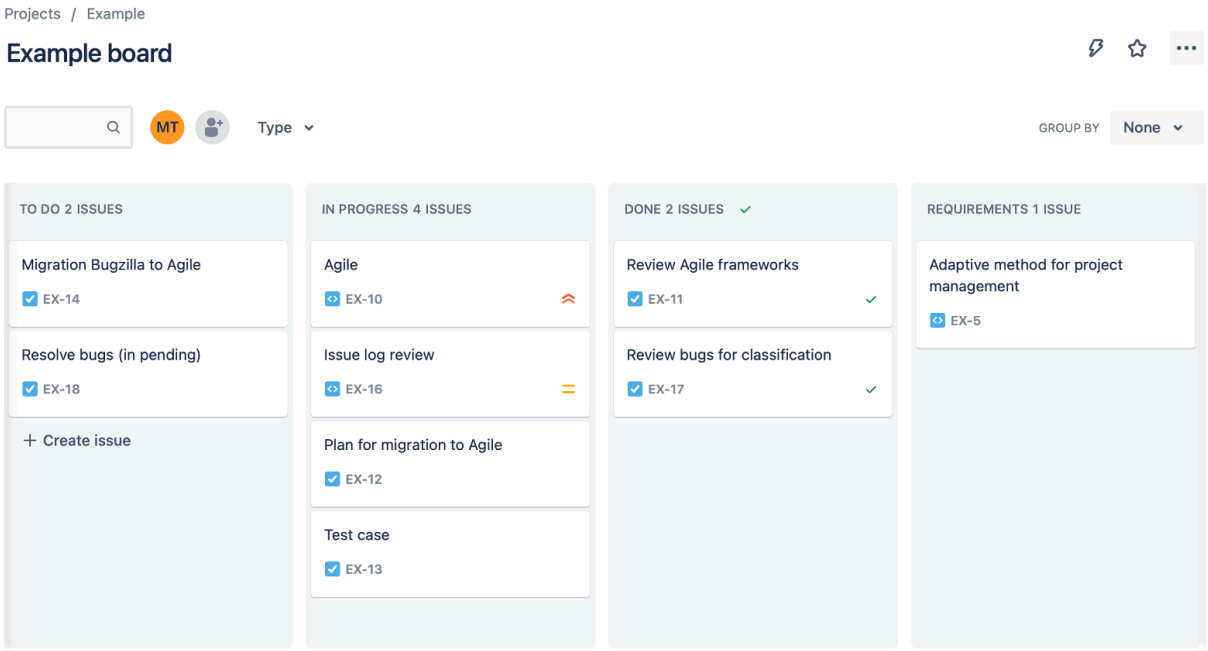

Figure 3.13 – Screenshot of the implementation of the example defined in Section 3.3.3 (Jira)

Finally, Figure 3.13 represents the implementation in Jira of the model example we presented in the previous section. The requirement, work items, and tasks created for the example are depicted in a Kanban board, which contains an additional column that we created to store the requirements being implemented.

3.5 Conclusion

As mentioned, an adaptive approach like Agile can overcome some of these inefficiencies of the standards development process, but it also brings its own challenges. In Agile, requirements are defined, analysed, and processed more frequently, and therefore, properly capturing and managing them becomes more time-consuming, and more complex. In Agile, requirements management becomes a bigger activity. Adopting Agile for standards development comes with an increased need for traceability and visibility of requirements, their elicitation, their management, and their implementation. In this chapter, we propose a new requirement elicitation model that serves as a foundation for requirements and decision traceability and visibility, by recording and leveraging project meetings minutes in a semi-formal way. We also present an implementation of this model through customisation of Jira, a bug and issue tracker used by standards development teams.

Standards support interoperability by facilitating the communication and collaboration between heterogeneous systems. Standards are a reliable solution to address data interoperability challenges. However, the complexity and length of their development process prevent them from being up to speed with the fast pace of the industry, and the quick and constant evolving nature of threats. To address the data interoperability challenge, we propose to improve the traditional information standard development process by 1) transitioning from predictive to adaptive project management and adopting Agile methods to shorten the development iterations and reduce the administrative overhead, and 2) introducing a model for better requirement elicitation as well as enhanced traceability and visibility for all the standards development stakeholders. Now, we are going to focus on the second challenge, data traceability.

CHALLENGE 2: DATA TRACEABILITY

CHAPTER 4: UNDERSTANDING CYBERSECURITY IN MANUFACTURING

Due to the growing diversity and severity of cyberattacks, organisations are investing more money in cybersecurity solutions to improve their security posture. Industry 4.0 has enabled industry manufacturing to become more digitalised and more connected, which also led to an increase in the complexity and volume of cybersecurity challenges. A solution to improve cybersecurity strategies and manage cyber risks is the adoption of a cybersecurity framework. This chapter focuses on cybersecurity and cybersecurity frameworks. Section 4.1 introduces cybersecurity and digital threats that are a growing concern for organisations. Section 4.2 presents cybersecurity challenges in the context of manufacturing. Section 4.3 introduces cybersecurity frameworks, a solution to protect organisations against these cyber threats.

4.1 Introduction to cybersecurity

Over the years, computer networks and information technologies have become an essential aspect of the functioning of society, economy, and critical infrastructures. Networks of machines connected to the internet are used to control and manage manufacturing processes, water supplies, the electric power grid, air traffic control systems, and stock markets, among others [166]. The growing reliance of governments, companies, and financial institutions on the Internet and other information technology, makes cyber-attacks more appealing and potentially more destructive. Cyber-attacks are defined as attempts by unauthorised people to access computer networks or systems with the intent of stealing, disrupting, damaging, destroying, or performing other illegal acts. Data and personal information are common targets of these attacks. A data breach is a security incident in which sensitive, confidential, or protected information is accessed without the knowledge or authorisation of its owner. According to [18], the average cost of a data breach was $4.24 million globally in 2021, and since 2015, the cost of a data breach has increased by 11.9%. These costs include the expenses of discovering and responding (i.e., detection and escalation, notification, and post breach response), as well as the cost of lost business (e.g., business disruption and revenue losses from system downtime, cost of lost customers, as well as reputation losses) [18]. The main causes of data breaches are malicious or criminal attacks, flaws in IT systems, and human error [167]. The majority of these malicious attacks are caused by either internal (e.g., employees abusing their access to the system and data) or external (e.g., use of stolen credentials or computing devices, social engineering, or hacking such as using malware or viruses, exploiting weak credentials or systems vulnerabilities) malicious actions [167].

Information technologies, digital data, and computer networks have become indispensable for businesses, governments, financials, and educational institutions. As a result, cybersecurity has become a major concern for all organisations, regardless of their size and sector. Besides, the rapid development of technologies is also making cybersecurity more challenging and essential. In 2013, the President of the United States of America, Barack Obama, signed an executive order designed to protect the nation's critical infrastructure against cyber risks [168]. Cyber threats have been declared as "among the gravest national security dangers to the United States [169], and are considered as a growing threat that affects citizens, government, and private companies. As mentioned, cyberattacks are costly to organisations and the global economy, and as a result, the cybersecurity market value is growing as companies and governments are investing a lot in cybersecurity to strengthen their security posture.

Cybersecurity refers to the technologies, processes, and best practices for protecting critical assets (e.g., hardware, software, networks, and systems) and sensitive data against any type of cyber-attack. [170] defined cybersecurity as "the organisation and collection of resources, processes, and structures used to protect cyberspace and cyberspace-enabled systems from occurrences that misalign de jure from de facto property rights." This definition represents the dynamic nature of cybersecurity and refers to the protection against cybersecurity incidents which is an event that misaligns actual property rights from perceived property rights, whether intentionally or accidentally, and whether predictable or unpredictable.

Among others, cybersecurity is key to maintaining and protecting the confidentiality, availability, and integrity of data. The CIA (Confidentiality, Integrity, and Availability) triad is a fundamental concept of information security [171]. Organisations broadly apply the CIA triad to all of their information security controls, processes, threats, and vulnerabilities. The three main principles of the CIA triad security model are:

- Confidentiality: it refers to the efforts of an organisation to protect its sensitive data from unauthorised access. It means ensuring only authorised individuals have access to specified data and preventing unauthorised individuals from gaining access to this data. Confidentiality can be breached through i) direct attacks (i.e., gaining illegal access to systems) such as escalation of system privileges or man-in-the-middle attacks, for example, ii) human error (e.g., weak passwords, shared user accounts, poor or absence of authentication systems) and iii) inadequate security measures. Confidentiality can be protected through several measures like robust authentication mechanisms, access controls, and data encryption;
- Integrity: it refers to the efforts of an organisation to ensure that data transmitted, processed, and stored has not been tampered with (deletion or modification), either accidentally or maliciously. In other words, it means ensuring that data is accurate, authentic, and trustworthy, and therefore can be

trusted. Integrity can be compromised in different ways through a change of system logs, the introduction of malicious code in databases, modification of configuration files, man-in-the-middle attacks, and human error for example. Integrity can be maintained with the help of preventive measures such as encryption, digital signatures and certificates, intrusion detection systems, strong authentication mechanisms, and version control, among others;

- Availability: it refers to the efforts of an organisation to ensure that authorised users get timely and reliable access to the organisation's assets (e.g., networks, systems, applications, and data) whenever they need them. Availability can be impacted by denial-of-service attacks, hardware or software failure, power failure, or human error. Availability can be ensured through backups, regular software patching, system upgrades, systems redundancy, or even comprehensive disaster recovery plans.

Today's society (i.e., individuals, businesses, governments) relies greatly on information technologies, and this reliance keeps growing. Besides, in modern organisations, the number of users, devices, systems, and programs, as well as the amount of data generated, most of it being sensitive or confidential, is continuously growing. Combining this with the improvements in cloud services and the Internet of Things (IoT), a multitude of new potential security vulnerabilities have emerged. The dangers of cybercrime are exacerbated by the increasing volume and sophistication of cyber-attacks. As a result, the importance of cybersecurity continues to rise.

Government, military, corporate, financial, and medical organisations collect, process, and store massive amounts of data, which makes cyber security critical. A considerable amount of this data can be considered sensitive information, such as intellectual property, financial data, personal information, or other types of data that unauthorised access or exposure could have serious consequences. Organisations constantly exchange sensitive data across systems and networks, and between various devices. Protecting this data and the systems that process and store it is primordial because cyber-attacks can have a direct impact on the confidentiality, integrity, and availability of organisations' data, as well as ruin their relationships with customers and even lead to serious legal consequences. Weak or inadequate cybersecurity posture can damage an organisation in a variety of ways such as economic costs (e.g., theft of intellectual property and corporate information, expenses of fixing damaged systems), reputational costs (e.g., loss of consumer trust, loss of customers, negative media coverage) and regulatory costs (e.g., regulatory fines or sanctions) [172,173]. The international research and advisory firm Gartner Inc. predicted that the global spending on information security and risk management technology and services would reach $150.4 billion in 2021 and also stated that organisations' investment in cybersecurity and information security is increasing in 2021, making it the top priority for new spending [174]. The need to protect

sensitive data from cyber threats is a major concern for all organisations as the number of cyber-attacks and data breaches in recent years keeps growing. Some examples of high-profile data breaches are:

- In 2013, Yahoo suffered a massive data breach that was at first estimated to have compromised 1 billion accounts, but that breach was more extensive than previously reported, and the number of accounts affected was later updated to 3 billion [175]. This breach is considered one of the largest data breaches in history.

- In 2017, Equifax suffered a major data breach that compromised the personal information of 147.9 million customers [176]. As a result of this breach, Equifax suffered reputational damage as well as numerous lawsuits.

Despite the efforts of cybersecurity professionals to reduce security gaps, attackers are constantly seeking new ways to avoid detection, bypass defence measures, and exploit new vulnerabilities. The most common cybersecurity threats are malware, ransomware, phishing, social engineering, insider threats, distributed denial-of-service (DDoS) attacks, and man-in-the-middle attacks. Cybersecurity threats are getting more diverse and sophisticated, by taking advantage of the changes in work habits (i.e., work-from-home environments and remote access tools) for example. The last few years were marked by an enhancement of the complexity of our cyber exchanges and strong acceleration in cyber threats due to the generalisation of teleworking and the massive use of communication and digital networks. The massive adoption of teleworking requires remote access to systems and data that can be sensitive, making them vulnerable and prone to cyberattacks. The inherent vulnerabilities associated with the significant adoption of remote work have led cybercriminals to increasingly target remote workers. According to [177], remote workers were responsible for security breaches in 19.8% of organisations.

The changing nature of security risks is one of the most difficult aspects of cybersecurity. The frequent changes and improvements in cyber threats are challenging for organisations that need to regularly update their cybersecurity strategies to protect against them as well as new potential vulnerabilities. This can be especially arduous for small and medium-sized organisations due to limited resources and insufficient cybersecurity expertise. In recent years, cyber threats have become a major problem for all businesses, regardless of their size and sector. According to the Verizon Business 2020 Data Breach Investigations Report, even though the gap between small and medium-sized companies, and large companies in terms of data breaches remains large (28% of data breaches involved small and medium-sized companies, while 72% involved large companies), the dividing line between small and large companies is reducing [178]. According to the Small Business Association (SBA), “small businesses are attractive targets because they have information that cybercriminals want, and they typically lack the security infrastructure

of larger businesses" [179]. Large businesses are generally better protected than small businesses and it is, therefore, easier for cybercriminals to infiltrate small businesses' systems and networks without being noticed. This can explain why cybercriminals, who aim to get the information they want using the quickest and easiest route, are targeting more and more suppliers of large companies, which are generally small and medium-sized companies. A BlueVoyant study states that 93% of companies surveyed have experienced a cybersecurity breach resulting from a weakness in their supply chain, and 97% of companies surveyed have been negatively impacted by a cybersecurity breach in their supply chain [180]. Supply chains are the perfect targets for cyber-attacks, as organisations do not properly monitor their vendors. However, significant supply chain attacks, like Solarwinds [181], in the last few years have exposed the extent of the threat against supply chains. The BlueVoyant study also reveals that only 13% of companies indicated that third-party cyber risk was not a priority, compared to 31% last year, which demonstrates a growing awareness of the problem, but continuous monitoring of their vendors remains low [180].

4.2 Cybersecurity in manufacturing

The fourth industrial revolution, also known as Industry 4.0, refers to the digital transformation of the business world. It also describes the current trend in automation of and data exchange in manufacturing processes. The EEF's 2018 Cybersecurity Report found that 91% of the manufacturers are investing in digital technology [182] due to its significant impact on the productivity and efficiency of manufacturing companies thanks to the digitisation of machines and the automation of processes.

Industry 4.0 has been driven by many technological advances that have been blurring the lines between information technology (IT) and operational technology (OT) to bring digital transformation to the manufacturing floor, as well as between the physical and the digital worlds to create interconnected systems in which machines, sensors, software, products, and people interact. These technologies enable the collection of data in real-time in order to optimise processes and make decisions faster.

The recent digitsation of manufacturing has increased the amount of data that organisations collect, process, and exchange. Manufacturing companies are constantly looking for ways to improve their productivity and profitability while reducing costs and optimising resources. The new technologies of Industry 4.0 generate large volumes of data, which helps to improve machine performance and product quality. Data has become a crucial tool for manufacturing.

This digital transformation brings many benefits to the manufacturing industry, but it also exposes these organisations to new risks. Manufacturing has become more automated, connected, and data-driven, making it a more accessible target for cyber criminals. Industry 4.0 technologies connect physical and virtual systems and facilitate the exchange of data across and between multiple systems in real-time. This

expands the vulnerability surface of an organisation and makes it more difficult to protect. Organisations are today exposed to a variety of cyber-attacks that can disrupt their entire business model. As mentioned, the manufacturing sector was particularly affected by cybercrime in 2020 and 2021, as it ranked respectively second and first most attacked sector [14]. In the context of Industry 4.0, cybersecurity plays a critical role in preventing enterprises from losing their competitiveness. [183].

Cyber-attacks can have serious consequences for businesses such as destruction of the entire information system, financial losses, damaged reputation, and legal ramifications, among others. In some cases, cyber-attacks can cause an interruption or even a complete shutdown of an organisation's operations, resulting in a loss of revenue for the company during the entire time of the attacks and sometimes even after. Furthermore, cyber-attacks can also damage an organisation's reputation, which can impact the trust of customers as well as the trust of third parties with whom the organisation works. The EEF report shows that 24% of manufacturers admitted that they have sustained financial or other business losses due to cyber-attack [182].

Organisations handle and rely on large volumes of data, making data one of their most vulnerable assets. Manufacturers rely on digital data for collaboration between different systems and decision-making. As attacks against data are multiplying, common threats against data are 1) ransomware attacks that prevent users from accessing their systems either by locking their computer or by encrypting their documents and files until a ransom is paid, and 2) data manipulation, which consists of discreetly modifying the data to corrupt data-driven processes or decisions.

Data manipulation is a type of cyber-attack aiming to compromise the integrity of data. Data manipulation attacks happen when a malicious actor gains unauthorised access to sensitive and private information with the intent of altering, manipulating, or changing it. A successful data manipulation attack can be highly damaging to an organisation. These attacks can be simple, such as changing the value of one or two data at times, but the repercussions for the organisation can be important. Such attacks affect the trust of an organisation in its own data, as well as the trust of the consumers in the organisation. Data manipulation attacks can be difficult and time-consuming to detect. They can also have a significant impact on an organisation, leading to the corruption of decisions and processes that rely on the compromised data, as well as the propagation of corrupted data within and across organisations. A common occurrence is void attack, in which tampering can result in structurally weaker parts and premature failure as well as functionally different parts and physical hijack. Such an attack would be difficult to detect and would not necessarily cause any warning during inspection or printing, if a minimal amount of data in the part file is tampered with and if the void is situated in a structurally important location.

The impact of data tampering varies depending on the data that has been compromised but can often cause serious and costly damage to organisations such as disruption of supply chains or physical tampering of goods. Preventing an attack aimed at data manipulation is difficult but critical. Organisations must learn to respond to such attacks. The longer it takes for an organisation to identify a data manipulation attack, the more damage it will suffer. Organisations produce and consume large volumes of data, and this data is exchanged among various systems, units, or organisations. However, in the case of data tampering, these exchanges also allow the propagation of compromised data within the organisations' systems, and sometimes across organisations, which can lead to serious consequences for an organisation and its partners (e.g., suppliers and customers) [184].

Cyber threats are constantly evolving and diversifying, requiring organisations to intensify their cybersecurity efforts. Organisations and their employees must implement cybersecurity best practices to limit risks and protect assets against cyber threats.

4.3 Introduction to cybersecurity frameworks

Cyber-attacks are growing in volume and sophistication, and in response, organisations need to implement practices and tools to protect their sensitive information. A 2020 IBM report highlights “a growing divide in data breach costs between organisations with more advanced security processes, like automation and formal incident response teams, and those with less advanced security postures in these areas” [185]. The ever-evolving nature of security risks alongside the ever-changing security threat landscape makes it really challenging for organisations to maintain an effective cybersecurity posture. Traditionally, organisations focus their efforts and resources on i) defending their systems against known threats, and sometimes only against the most well-known threats, leaving lesser-known threats undefended, and ii) protecting their most critical assets, sometimes to the detriment of other components that remain with limited protection or even unprotected. Both of these strategies are insufficient because i) attackers can develop new attacks or modify existing attacks, and then the organisations would be vulnerable, and ii) attackers could then target less critical assets made more vulnerable by a lack of protection in order to gain access to the organisation's networks and systems. Threats evolve and change quickly, and, in this context, organisations need to apply a more proactive and flexible approach in their risk management efforts.

New threats regularly emerge and adapt to counter the cyber defences put in place by organisations. As a result, it is necessary to have solutions that remain up to date with the constant changes and progress made in terms of cyber threats. Each organisation should ensure that its cybersecurity policy and the education of these employees on cybersecurity are regularly updated to avoid potential vulnerabilities or behaviour that could undermine the security of the organisation. In fact, organisations, regardless of the

size, should ensure that their employees understand cybersecurity threats and how to mitigate them through regular cybersecurity awareness training courses and the use of a framework aimed at reducing the risks of data breaches. As mentioned previously, organisations are investing more money in cybersecurity solutions to strengthen the security of their networks and systems, as the diversity and severity of cyberattacks grow.

Furthermore, governments around the world have become aware of the cyberthreats hanging over their citizens' data and personal information, and thus have adopted comprehensive laws to ensure data protection and privacy such as the European Union's General Data Protection Regulation (GDPR), the Personal Information Protection and Electronic Documents Act (PIPEDA) in Canada, Brazil's General Personal Data Protection Law (LGPD), the United States' Health Insurance Portability and Accountability Act (HIPAA) or the California Consumer Privacy Act (CCPA) for example.

In addition to those laws, many solutions can assist organisations in protecting their assets against cyber threats. One of these solutions is to adopt a cybersecurity framework. Managing the cybersecurity of an organisation can be complex and adopting a cybersecurity framework can help facilitate it. A framework provides a structure and a methodology for organisations to protect and monitor their assets, plan their responses, and educate their employees on best practices. More specifically, a cybersecurity framework provides a set of standards, recommendations, guidelines, and best practices to enhance cybersecurity strategies and manage cyber risks. The requirements and practices described in a cybersecurity framework help an organisation's security team define reliable and flexible procedures to assess, monitor, and mitigate cybersecurity risks. In some cases, organisations may be required to implement cybersecurity frameworks in order to demonstrate compliance with industry or international cybersecurity regulations. For example, organisations that would like to accept, store, transmit or process credit card information must prove that they comply with the Payment Card Industry Data Security Standards (PCI DSS) framework.

There are many cybersecurity frameworks, and these frameworks can be split into several categories. At an RSA conference in 2019, Frank Kim, founder of ThinkSec and former CISO at the SANS Institute, explained the various framework types. He divided them into three categories: control frameworks, program frameworks, and risk frameworks [186]. The control frameworks represent the foundations of all security programs. They specify the controls and activities that an organisation must implement in order to be protected from cyber-attacks. Program frameworks offer a higher-level view of the overall security activities that an organisation is doing. This type of framework facilitates communication with business leaders to make them better understand what is being done from an overall security perspective. Finally, risk frameworks provide a way to better understand how to prioritise security activities and determine what needs to be done/protected first. These three types of frameworks can be used together, they are not

mutually exclusive and can complement one another. The more a company grows, the more its security program will expand and integrate more and more elements from control, program, and risk frameworks.

Cybersecurity frameworks are the main tool organisations can use to learn about attacks and response strategies. Cybersecurity frameworks are generally quite generic and unopinionated, providing high-level guidelines for managing risk (i.e., prepare, identify, and respond to) against cyber threats. Organisations face numerous cybersecurity challenges. Organisations can use cybersecurity frameworks to bring new security layers and improve their security programs. As a result, government agencies and private companies have developed many cybersecurity frameworks to help detect and reduce security gaps.

4.4 Conclusion

Businesses and governments rely more and more on information technologies, computer networks, and digital data. As a result, cybersecurity has become a top priority for all organisations, regardless of their size and sector. As manufacturing has become more digitalised, connected, and data-driven, it has also become more vulnerable to digital threats, making data security a major challenge for Industry 4.0. Among the various cyber-attacks, data manipulation has been confirmed to be a serious threat to organisations with significant consequences. As cyber-attacks are growing in complexity and sophistication, organisations are investing more and more money in cybersecurity solutions to strengthen their security posture. Adopting a cybersecurity framework is one solution that allows organisations to enhance their cybersecurity strategies and protect their assets against cyber threats. Many cybersecurity frameworks exist and in the next chapter, we will review and evaluate a set of comprehensive cybersecurity frameworks in an effort to find the best candidate to support our needs for addressing data manipulation risk.

CHAPTER 5: A CYBERSECURITY FRAMEWORKS REVIEW

The digital transformation of manufacturing comes with a growing number of cybersecurity challenges. Data manipulation is one of these challenges and it entails discreetly modifying data to corrupt organisational data-driven processes or decisions. Cybersecurity frameworks are a solution to help mitigate cyber risks, but they are often generic and unopinionated to attract a larger audience, leaving their users to customise and implement them. This chapter focuses on identifying a generic yet opinionated framework that can be customised to support our goal of addressing data manipulation risks. Section 5.1 details the selection process of this framework. Section 5.2 presents the benefits, adoption, and limitations of the chosen framework as well as its customisation mechanisms.

5.1 Review process

5.1.1 Initial selection

Many cybersecurity frameworks exist and, we review and evaluate a set of 36 cybersecurity frameworks to find the best candidates to address Industry 4.0 cyber challenges mentioned in the previous chapter. The initial set includes the most comprehensive frameworks developed and adopted by standards development organisations (SDOs), government agencies, and industrial consortiums. The initial set was screened in four iterations using inclusion criteria, one after the other, and at each iteration, a new set was created by removing the frameworks that did not meet the criterion. The criteria used to evaluate these frameworks are:

1. **cybersecurity compliance framework:** Cybersecurity compliance frameworks and regulatory compliance frameworks are two types of compliance frameworks. Both types of frameworks include a set of guidelines, recommendations, procedures, and tools that assist organisations in protecting customers across the world. The fundamental difference between these two types of frameworks is that cybersecurity compliance frameworks mainly focus on data while regulatory compliance frameworks are more consumer-oriented [187]. Furthermore, regulatory compliance frameworks are mandatory practices, and failing to comply with them can have serious consequences for organisations such as financial penalties or inability to conduct business. On the other hand, cybersecurity compliance frameworks are recommendations for establishing and maintaining a security posture as well as managing the evolving nature of data asset protection [187]. This is why the set was screened to remove all non-cybersecurity compliance frameworks.

2. **comprehensive framework:** The primary objective of our work focuses on data security in the context of Industry 4.0. Industry 4.0 encompasses many different manufacturing domains, and our solution needs to be applicable to all of these domains while also being able to manage the specifics of several domains at the same time. As a result, we choose to keep comprehensive cybersecurity frameworks that can be used by any sort of organisation and remove industry- or domain-specific frameworks that aim to achieve a specific objective [188].
3. **standard-based framework:** Cybersecurity frameworks are typically defined as a set of guidelines, information requirements, tools, and best practices for managing risk and protecting organisations against cyber-attacks. The use of information standards to document these elements provides multiple benefits:
    a. It facilitates the human consumption of the frameworks by leveraging agreed-upon graphical notations (e.g., process models in BPM or information models in UML) that are well-documented and traditionally known to subject matter experts.
    b. It facilitates the computer consumption of the frameworks by leveraging data representation and exchange formats (e.g., XML, XMI) that are widely supported by open-source and commercial application programming interfaces (APIs).
    c. It facilitates the usage of the frameworks by relying on documentation formats that are widely supported by open-source and commercial tools such as UML designers or ontology editors.
    d. It facilitates collaboration and interoperability by providing dictionaries of common, widely accepted, and well-defined key concepts (e.g., existing ISO process definitions or taxonomies).
4. **framework with implementation guidelines:** Implementation is the process of transforming a strategic plan into action to achieve its objectives and produce deliverables. Implementation guidelines include, but are not limited to, software tools, IT infrastructure, audit processes, or data formats, to use in order to achieve strategic goals and compliance with the framework. A framework with implementation guidelines will facilitate and decrease the cost of its adoption by removing the overhead associated with designing, selecting, or developing the tools, infrastructure, or processes required.

The initial set was screened and a final set of four frameworks was reached (see Appendix). The four frameworks are ISO 27001, ISA/IEC 62443, NIST CSF, and SCAP.

5.1.2 Presentation of the final set

5.1.2.1 International Organization for Standardization (ISO) 27001

ISO/IEC 27001 was created collaboratively by the International Organisation for Standardization and the International Electrotechnical Commission (IEC). It provides best practices and information security controls to manage information security risks. ISO 27001 also outlines requirements for implementing, maintaining, and improving an Information Security Management System (ISMS) over time to ensure that it aligns with the needs and objectives of the organisation [189]. ISO 27001 defines a framework to help organisations, of any size or industry, protect their sensitive information in a systematic and cost-effective manner [190].

The standard is divided into two parts. The first and main part consists of 11 clauses that specify requirements that must be met by organisations that want to be compliant with the standard. While the second part, Annex A, defines 114 controls that are divided into 14 domains and support the clauses and requirements stated in the first part. Figure 5.1 represents the clauses of the first part and the domains of Annex A.

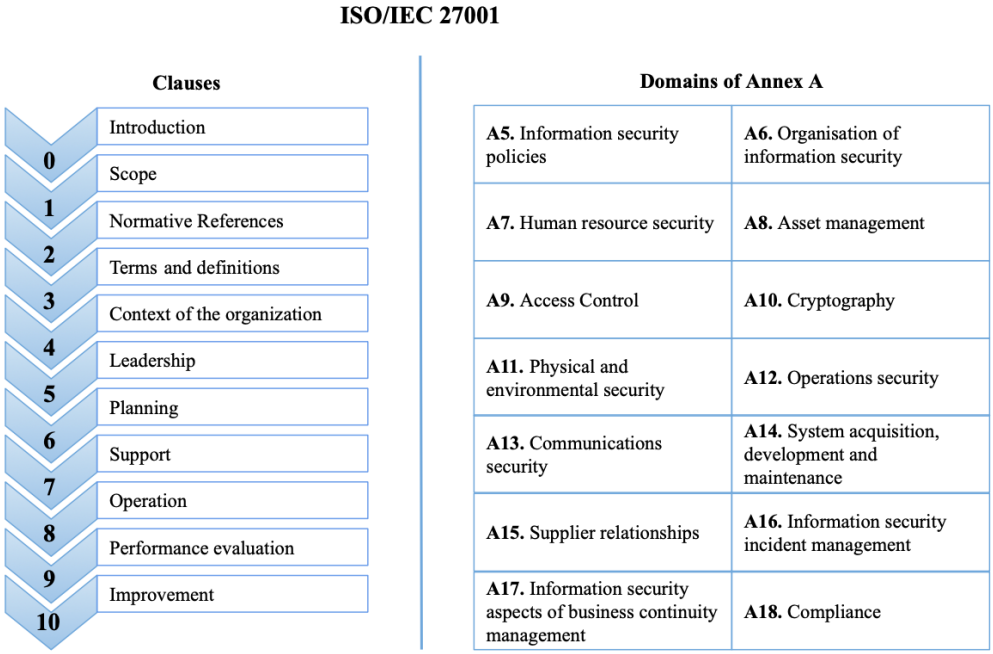

Figure 5.1 – ISO 27001 components

The 11 clauses of ISO 27001 are [191]:

- Clauses 0 to 3 serve as an introduction to the standard.
- Clause 4: Context of the organisation - it mandates an organisation to define the context of an ISMS.

- Clause 5: Leadership - it identifies the top management responsibilities for the ISMS implementation.
- Clause 6: Planning - it encompasses the actions an organisation must take to address information security risks.
- Clause 7: Support - it focuses on resources and how an organisation must secure the right resources, the right people, and the right infrastructure to implement, maintain and improve the ISMS.
- Clause 8: Operation - it concerns the execution of the plans and processes defined to achieve the information security objectives.
- Clause 9: Performance evaluation - it provides requirements to help organisations monitor, measure, and assess the performance of the ISMS.
- Clause 10: Improvement - it includes corrective action requirements an organisation should take to ensure continual improvement of the ISMS.

The 14 domains listed in Annex A cover six security areas: company security policy, asset management, physical and environmental security, access control, incident management, and regulatory compliance.

5.1.2.2 International Society of Automation - ISA/IEC 62443

The International Society of Automation (ISA) and the International Electrotechnical Commission (IEC) collaborated to develop the 62443 series. ISA/IEC 62443 provides a series of requirements and methods for addressing and mitigating security risks and vulnerabilities in industrial automation and control systems (IACSs) [192]. This series aims to improve the safety, reliability, integrity, and security of IACS by implementing a risk-based, methodical and comprehensive process throughout their lifecycle. It covers both technical and process-related features of automation and control systems cybersecurity.

These requirements were originally developed for the industrial process industry, but they have now been expanded to include building automation, medical devices, and transportation industries. The ISA/IEC 62443 series can benefit asset owners, product suppliers, and service providers.

The ISA/IEC 62443 series standards and technical reports are organised into four groups with 14 implementation documents, as shown in Figure 5.2. Each group corresponds to a primary focus and target audience [193]:

- General - it includes four implementation documents that cover subjects that are common to the entire series;
- Policies and Procedures - it contains five implementation documents that detail the policies and procedures associated with IACS security;

- System - it consists of three implementation documents that focus on system-level requirements;
- Component - it includes two implementation documents that address requirements related to the development of IACS products.

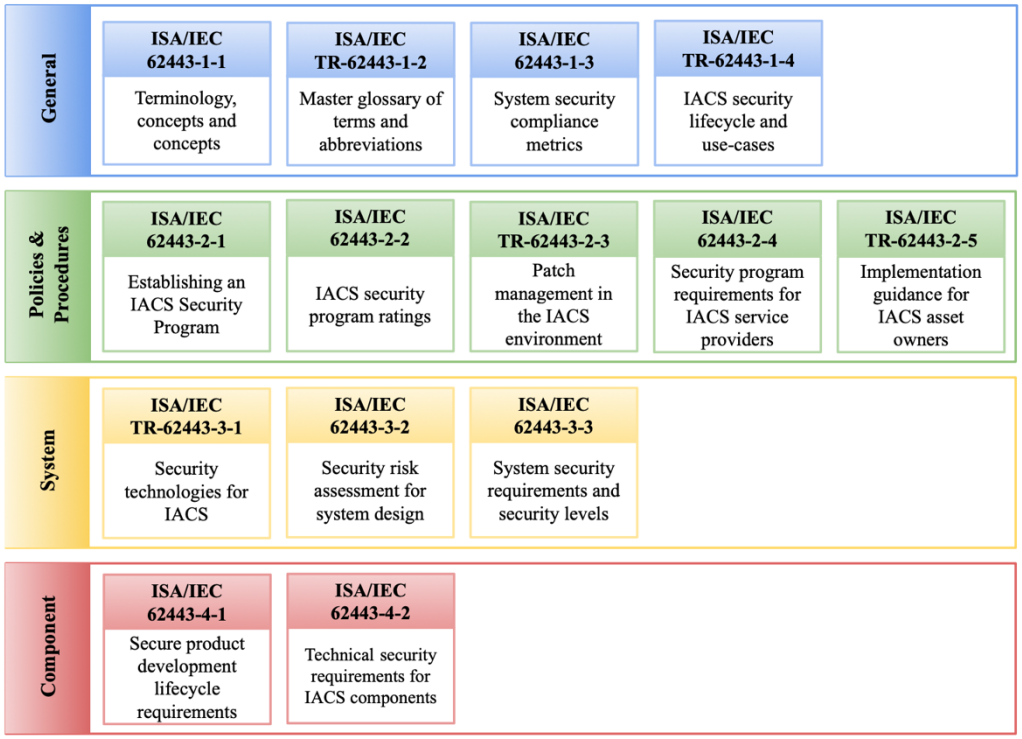

Figure 5.2 – ISA/IEC 62443 Standards series

5.1.2.3 National Institute of Standards and Technology (NIST) Cybersecurity Framework (CSF)

Following the issuance of Executive Order 13636, "Improving Critical Infrastructure Cybersecurity" in 2013, NIST developed the Framework for Improving Critical Infrastructure Cybersecurity, also known as the NIST Cybersecurity Framework. As a result of this Order, NIST had the responsibility to collaborate with stakeholders, both from industry and government, to develop a framework for mitigating cyber risks to critical infrastructure based on existing standards, guidelines, and successful practices [194]. The Framework offers a prioritised, flexible, repeatable, and cost-effective approach to managing cybersecurity risk in critical infrastructure environments [195]. Furthermore, the Framework also explicitly acknowledges that organisations have different cybersecurity risk management needs, which necessitates different types and levels of cybersecurity investments [196]. The Framework is a cybersecurity risk management solution that enables technical innovation, adapts to any organisation, regardless of its size or sector, and evolves

with technological advancements and business requirements [195]. The Framework can be used to create a new cybersecurity program or as a tool to identify and improve gaps in existing cybersecurity approaches.

The NIST Framework is a risk-based approach and is composed of three main parts: the Framework Core, the Framework Implementation Tiers, and the Framework Profiles. Each of these components emphasises the connection between business drivers and cybersecurity activities. According to NIST [195]:

- The Framework Core provides a set of cybersecurity activities and desired outcomes using a common language. It helps the communication of cybersecurity activities throughout the organisation, from the executive level to the operation level. The Core is composed of five Functions - Identity, Protect, Detect, Respond and Recover – and describes key Categories and Subcategories for each of these Functions. The Functions and Categories of the Framework Core are depicted in Figure 5.3. The Functions help organisations manage cybersecurity threats. Each Subcategory has been mapped to controls of international industry cybersecurity standards.

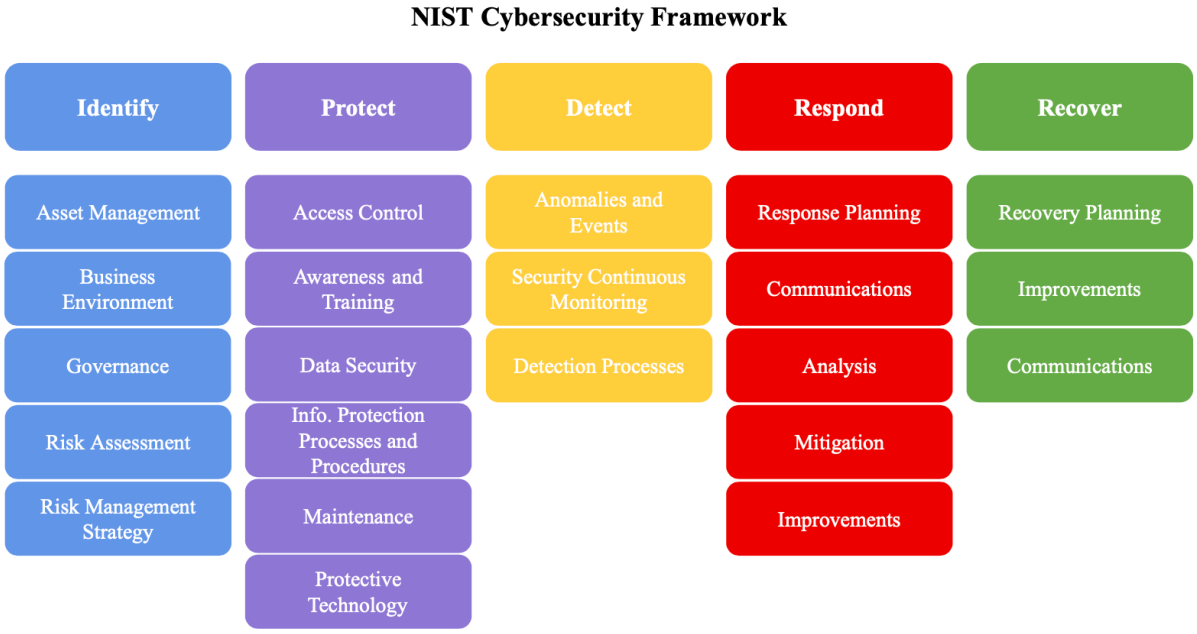

Figure 5.3 – NIST Framework Core

- The Framework Implementation Tiers provide context on how an organisation views cybersecurity risk and the processes in place to address it. The Tiers provides guidelines that help organisations decide the appropriate level of rigour for their cybersecurity program. The Tiers define implementations ranging from Partial (Tier 1) to Adaptative (Tier 4) to characterise an organisation's practices.

- The Framework Profiles describe the outcomes an organisation has chosen from the Framework Categories and Subcategories. Profiles represent the alignment of a “Current” Profile (i.e., current organisational requirements and objectives, risk appetite, and resources of an organisation) with a “Target” Profile (i.e., the Framework Core’s outcomes selected by the organisation). Organisations commonly use Profiles to identify opportunities to strengthen their cybersecurity policies.

5.1.2.4 Security Content Automation Protocol (SCAP)

The Security Content Automation Protocol (SCAP) is a set of specifications based on specific standards that allow for automated vulnerability checking, monitoring, and management as well as security policy compliance checking [197]. SCAP standardises the processes of communication (formats and nomenclature) of security products and tools. It aims to enable organisations to measure, express, and assess security data using standardised criteria.

Automated solutions using SCAP specifications make it easier to continuously monitor the security compliance status of a wide range of IT systems. Using standardised, automated practices for system security management can help organisations function more efficiently in complex, interconnected environments while reducing costs.

SCAP is composed of numerous components as represented in Figure 5.4:

- Common Vulnerabilities and Exposures (CVE) provides unique identifiers to publicly known security vulnerabilities and exposures [198].
- Common Platform Enumeration (CPE) is a standardised method for describing and identifying classes of applications, operating systems, and hardware devices that can be found within an organisation's IT assets [199].
- Common Configuration Enumeration (CCE) provides unique identifiers to security system configuration issues in order to facilitate rapid and accurate correlation of configuration data across multiple information sources and tools [199].
- Open Vulnerability and Assessment Language (OVAL) is a common language for declaring the state of a system, describing security vulnerabilities or desired system configurations [198].
- Extensible Configuration Checklist Description Format (XCCDF) is an XML-based language for expressing security checklists, benchmarks, configuration documentation, and other related documents [199].
- Common Vulnerability Scoring System (CVSS) is a method to describe the principal characteristics of a vulnerability and measure its impact and severity. This score can be used by organisations to accurately assess and prioritise their vulnerability management activities [200].

- Open Checklist Interactive Language (OCIL) is a common language to communicate a series of questions to be given to a user and to interpret the responses to these questions [199].
- Asset Identification (AID) provides methods to uniquely identify assets based on known identifiers and/or known information about the assets, as well as guidelines on how to use this identification [199].
- Asset Reporting Format (ARF) is a data model that outlines the transport format of information about assets, and the relationships between assets and reports [199].
- Common Configuration Scoring System (CCSS) provides a way to quantify the severity of software security configuration issues [201].
- Trust Model for Security Automation Data (TMSAD) is a data model for ensuring trust that can be applied to specifications within the security automation domain [199].
- Software Identification (SWID) Tagging allows organisations to track the software installed on their devices in a transparent manner [202].

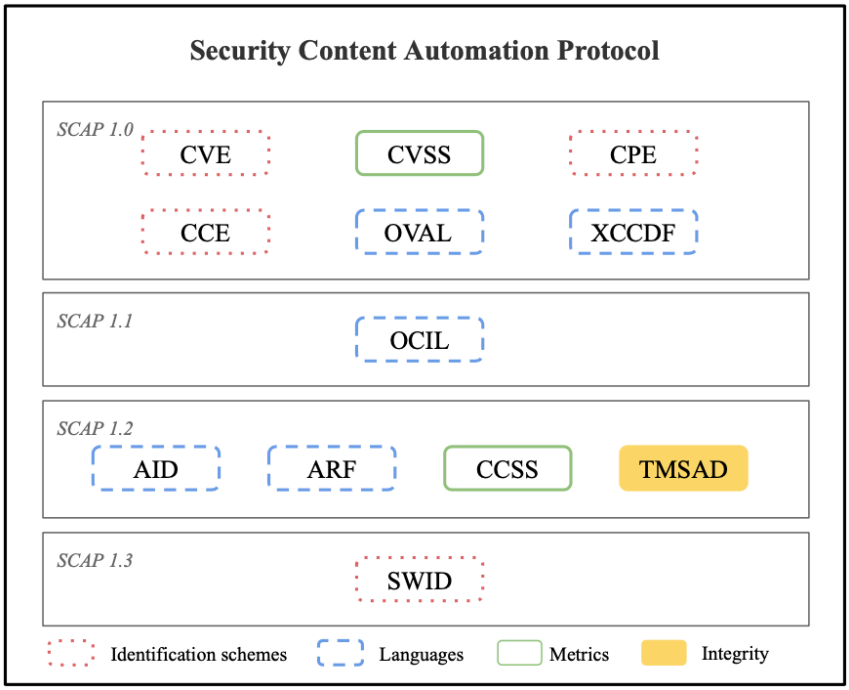

Figure 5.4 – SCAP Components

5.1.3 Final selection

A final selection is made to determine which one of the four frameworks presented above would best meet our needs in terms of protecting data integrity against data manipulation attacks. These attacks aim to compromise the integrity of data by subtly tampering it. In the event of such an attack, it is important to be

able to trace the data flows in order to precisely determine which systems have consumed or propagated tampered data and quarantine them, without having to completely shut down all systems [184].

Unfortunately, none of these four cybersecurity frameworks have a native support for data integrity and data traceability and therefore, none meets our key requirements. As a result, an additional criterion was added to evaluate customisability (and extensibility). This criterion helps identify the framework that is the most easily customisable to support our objective. The results of this final selection are summarised in Table 5.1.

Table 5.1 - Final selection process.

Frameworks	Integrity for traceability	Customisable
ISO 27001	X [(1)]	✓ [(4)]
ISA 62443	X	X
NIST CSF	X [(2)]	✓ [(5)]
SCAP	X [(3)]	✓ [(6)]

[(1)] The goal of ISO 27001 is to protect the confidentiality, integrity, and availability of information.

[(2)] Two implementations (NIST SP1800-25: Data Integrity: Identifying and Protecting Assets Against Ransomware and Other Destructive Events [203], and NIST SP1800-26: Data Integrity: Detecting and Responding to Ransomware and Other Destructive Events [204]) demonstrate the application of the NIST Cybersecurity Framework to data integrity, but with a focus on ransomware and destructive events.

[(3)] One of the components of SCAP is the Trust Model for Security Automation Data (TMSAD) which consists of digitally signing source data streams to ensure the integrity and trustworthiness of the content [205].

[(4)] The implementation of ISO 27001 controls can be customised to suit an organisation's environment.

[(5)] The NIST CSF was designed to be customisable [206] One of the components of the Framework, the Framework Profile, is meant to be used as a customisation of the Core to meet the needs and practices of a given sector, organisation, or overall objective [207].

[(6)] Certain parts of SCAP, such as XCCDF and OVAL-based checklists, are customisable. A checklist defines instructions for securely configuring an IT product for a specific environment or verifying that it has already been securely configured. Checklists can be customised to meet the needs and characteristics of an organisation [208].

As mentioned, none of the four frameworks are specific to data manipulation, therefore it is necessary to choose a framework that can be customised in order to meet our own requirements on data manipulation. As presented in Table 5,1, three frameworks (ISO 27001, NIST CSF and SCAP) meet our customisation criterion. Our final choice is the NIST Cybersecurity Framework for the following reasons:

- SCAP comprises multiple components, but only some of them (e.g., XCCDF and OVAL-based checklists) are customisable. Moreover, the scope of SCAP is not as important as that of the NIST CSF. SCAP focuses on vulnerability management while the NIST CSF provides comprehensive guidance to build strong cybersecurity programs.
- While both NIST CSF and ISO 27001 have similar scopes and are customisable, 1) the NIST Cybersecurity Framework is the framework the most widely used, according to the 2019 SANS OT/ICS Cyber-security Survey [209]; 2) it was intentionally designed to be customisable thanks to its tailoring mechanism known as "Profile"; 3) the NIST CSF provides a risk-oriented approach, that focuses on identifying and resolving security gaps [65]; 4) the NIST CSF offers some flexibility in implementation for organisations [210,211]; 5) ISO 27001 is often perceived as challenging to navigate [212,213], while the NIST CSF is considered easy to use [211,214] as it refers to a set of relevant standards [210] and provides some level of opinionation: the Framework defines several Categories and Subcategories for each Function, and each Subcategory includes Informative References to existing standards, frameworks, guidelines, and practices. These references map the NIST CSF to other frameworks and help organisations implement it by providing additional information on how to achieve the desired outcomes described in the associated Subcategory. These references include ISO 27001 and ISA 62443.

5.2 Final solution: NIST Cybersecurity Framework

5.2.1 Benefits

The NIST Cybersecurity Framework was created with the intention of being a voluntary framework for improving cybersecurity and reducing cyber risk for critical infrastructure in the United States [215]. The Framework has achieved this purpose and has also become a platform for discussing cybersecurity. It gives organisations a mechanism to organise and lead the conversation about security goals and enhancements [216,217]. More generally, the NIST Cybersecurity Framework has become a major element of the national dialogue about cybersecurity, sparking debate about the measures needed to enhance cybersecurity on a large scale in the US and even internationally [216,217]. Even though it was originally intended for critical infrastructure sectors, the Framework can benefit any type of organisation and help

them manage their cyber risks [218,219,220]. It has already been used in discussions in the United States to encourage rigor towards cybersecurity in industries where such obligations are not yet mandatory [221].

The NIST Framework provides a common language for discussing cybersecurity challenges within an organisation, between the different business units, and between organisations [196,218,222,223]. Thus, with the help of this common language, the framework helps organisations describe cyber risks and communicate their current cybersecurity strategies as well as the improvements they wish to make to the various stakeholders (both inside and outside the organisation) [196,222,223]. This common language also enables organisations' leaders/executives to participate in identifying and improving their cybersecurity practices.

The NIST Framework offers organisations implementing it, regardless of their size and sector, a flexible and adaptable approach [196,224,225] that can be integrated with existing security programs. It provides a comprehensive overview of how to manage cybersecurity risks while being technology-neutral [224,226]. Organisations decide themselves on the details of its implementation (i.e., products, services…). In addition to its flexibility, the Framework was conceived to evolve with changing cybersecurity threats, processes, and technologies [195,227].

Based on the most recognised practices of several other industry standards, the NIST Framework is used by organisations to comply with other regulations, but also by smaller companies to protect their data and demonstrate their security readiness. The NIST Cybersecurity Framework is one of the most widely accepted approaches to cybersecurity risk management among organisations across all US industries, according to [196,228]. It is widely adopted by both critical and non-critical infrastructure organisations [229]. In 2015, it was estimated that 30% of US businesses have adopted the framework, and this number was expected to increase in the following years [230,231]. According to a 2018 study, the NIST Cybersecurity Framework was the most popular framework in healthcare with 57.9% of respondents saying they used it in their organisation [232]. Moreover, the majority of IT security professionals regard it as security best practice [221,227].

Table 5.2 – Summary of the NIST Cybersecurity Framework's benefits

Benefits	References
NIST CSF provides a mechanism to organise and drive conversations about security goals	[216,217]
NIST CSF has become major part of national conversations about cybersecurity	[216,217]
NIST CSF can benefit any kind of organisation, not just critical infrastructure	[218,219,220]

NIST CSF provides a common language to help the communication of cybersecurity challenges and approaches	[196,218,222,223]
It provides a flexible and adaptable approach	[196,224,225]
It is one of the most widely accepted approaches to cybersecurity risk management	[196,228]
It is considered a security best practice.	[221,227]

5.2.2 Adoption

The NIST Cybersecurity Framework takes up the challenge of adapting to different sectors, industries, and countries. It does not use specific standards, instead, it provides a list of potential international standards that could be used to satisfy its cybersecurity controls. It can be used in different fields such as critical infrastructure, government or private sector, as well as in different countries.

The NIST Framework has been adopted by many businesses, universities, and governments as part of their cybersecurity strategy. American companies including JP Morgan Chase, Microsoft, Boeing, and Intel use this framework, as do foreign companies like Bank of England, Nippon Telegraph and Telephone Corporation, and the Ontario Energy Board [233]. The University of Chicago, the University of Pittsburgh, and the University of Kansas Medical Center have implemented the NIST Framework to define their cybersecurity programs [234]. Likewise, major standards developing organisations such as ISO and ISACA have also incorporated the framework into their strategies. ISO has developed ISO/IEC TR 27103:2018 which provides guidance on how to use the NIST CSF to leverage existing standards in a cybersecurity framework [235,236].

Although the NIST Framework was developed by the US government to secure critical infrastructure sectors, it is used by governments and by both private and public companies, inside and outside of the United States. Furthermore, multiple countries have already adopted this framework for their cybersecurity strategy, with some even incorporating it into their national legislation [236,237]. To facilitate and encourage countries to use the Framework, it has been translated into several languages [238]. Among the countries that have adopted the CSF or aligned with it in their cybersecurity approaches are Bermuda, Israel, Italy, Japan, the UK, Scotland, Switzerland, and Uruguay [237].

In 2015, the Japanese industry decided to launch the Cross-Sector Forum with major companies from all the critical infrastructure sectors including chemical, financial, manufacturing, media, and transportation. These companies created “a good ecosystem to educate, recruit, retain, and train

cybersecurity professionals in collaboration with academia and government" [239,240,241]. The NIST Cybersecurity Framework was used by the Forum as a global standard for describing cybersecurity missions and terminologies that could be applied by any sector [239,240,241]. The NIST Framework provides a comprehensive view of cybersecurity talents organisations needs and cybersecurity missions they should pursue. Japan was the first country to translate the NIST Cybersecurity Framework. As a result, the NIST Framework serves as a common language for industry sectors and facilitates discussions about global cybersecurity practices between Forum members and their subsidiaries outside Japan [239]. In addition, the NIST Framework has been mapped to other information security standards such as COBIT and ISO/IEC 27001 which is the most commonly used information security standard in Japan, allowing the Forum to justify adopting the NIST Framework for their cybersecurity activities [239,241].

In 2015, the Research Center for Cyber Intelligence and Information Security (CIS) at Sapienza University of Rome released the Italian National Framework for Cybersecurity [242], which is the result of a collaboration between academia, public entities, and private companies. This framework, which is largely derived from the NIST Cybersecurity Framework, provides an approach to manage cybersecurity for public and private organisations of any size, but with a primary focus on small and medium-sized businesses with limited IT expertise [243,244].

The Uruguayan Cybersecurity Framework (MCU) was introduced in 2016 by the Agency Electronic Government and Information Society and Knowledge (AGESIC), within the Government of Uruguay. The MCU is based on the NIST Cybersecurity Framework and strives to improve cybersecurity and protect information security against cyber threats [245,246].

In 2015, the Australian Securities and Investments Commission (ASIC) released Report 429: Cyber resilience: Health check in which they advised organisations to use the NIST Cybersecurity Framework to identify and mitigate cyber risks or to assess their current cyber risk management practices [243,247]. In 2018, ASIC published Report 651: Cyber resilience of firms in Australia's financial markets: 2018–19. This report outlines key trends based on self-assessment surveys completed by financial markets firms, in which participants were asked to evaluate their cyber resilience against the NIST Cybersecurity Framework and identify existing good practices and opportunities for improvement [237,248].

Table 5.3 – Summary of the NIST Cybersecurity Framework's adoption

Adoption	References
American and foreign companies use the NIST CSF	[233]
American universities implemented the NIST CSF into their cybersecurity programs	[234]

Standards developing organisations have incorporated the framework into their strategies	[235,236]
Japanese Cross-Sector Forum	[239,240,241]
Italian National Cybersecurity Framework	[242,243,244]
Uruguayan Cybersecurity Framework	[245,246]
Australian Securities and Investments Commission Report 429 and Report 651	[237,243,247,248]

5.2.3 Limitations

As previously stated, the NIST Cybersecurity Framework is widely accepted and is considered a great guideline for creating or transforming an organisation's cybersecurity strategy. However, the NIST Framework has some limitations. First of all, the NIST Framework is complex [249]. The Framework Core is composed of five Functions, which consist of 23 categories and over a hundred subcategories, and each subcategory includes references to other frameworks such as COBIT, ISO, and ISA [195]. The framework can be difficult to comprehend and implement because these subcategories cover a broad range of topics and outcomes [249,250].

Implementing all of the activities listed in the NIST CSF's subcategories would strengthen an organisation's cybersecurity posture [251]. However, in addition to being complex, implementing the NIST CSF can quickly become time- and resource-consuming [252,253]. As a result, many organisations that already use the NIST CSF only implement some of the controls recommended in the framework and a majority of organisations planning to adopt the NIST Framework claimed that they would only implement certain controls [254,255]. The framework was designed to be scalable and can therefore be implemented gradually [206,256]. Although implementing the NIST CSF in its entirety is optimal, even implementing a subset of the NIST CSF activities correctly provides a great foundation for developing a cybersecurity program, especially for organisations with limited resources.

The needs of organisations are different. The NIST CSF will be used by some organisations as a foundation for their security strategy, while others will utilise it to complement or improve their existing programs. The framework is not one-size-fits-all, which means that there is not only one possible way to implement it and that it depends on the organisation (size, resources, threats, security posture) [195,257]. In addition, the framework can be difficult to navigate given the hundreds of activities it provides, knowing that the NIST does not define any level of priority for the subcategories. Organisations should decide for

themselves the order of implementing controls based on their needs [216]. Organisations should establish a prioritisation mechanism to effectively navigate the entire NIST CSF planning and implementation process. The Italian National Cybersecurity Framework, which is based on the NIST Cybersecurity Framework, offers a priority scale to help organisations classify the subcategories according to their needs [251]. Furthermore, the NIST Cybersecurity Framework stated that organisations should prioritise activities and investments to optimise the impact of each dollar spent [195] but provided no further specifics or guidelines on how investments should be prioritised. [196] proposes a model for integrating a cost-benefit analysis into the NIST Cybersecurity Framework in order to help organisations determine the appropriate Implementation Tier level they need.

As mentioned, the NIST Cybersecurity Framework was initially created to help critical infrastructure organisations manage cybersecurity risks. Any organisation can utilise the framework to mitigate and effectively manage cybersecurity risk. Implementing the NIST Framework would therefore reduce risks for organisations. However, one of the main limitations of the NIST CSF is that it lacks an analytic capability to quantify cyber risks [227,258,259]. It refers to best-practices recommendations and other NIST publications to help organisations measure risk themselves in a way that is appropriate for their resources and security strategy [252]. Several solutions have been proposed to help organisations assess cyber risks in a way that complements the NIST CSF or is consistent with these principles [258,259,260].

Table 5.4 – Summary of the NIST Cybersecurity Framework's limitations

Limitations	References
NIST CSF can be complicated to comprehend and implement	[249,250]
NIST CSF can be time- and resource-consuming	[252,253]
Organisations implement only some controls recommended in the framework, not all of them	[254,255]
It does not provide a level of priority for the implementation of the subcategories	[216]
NIST CSF does not provide an analytic capability to quantify cyber risks.	[227,258,259]

5.2.4 Customisation mechanism

The NIST Cybersecurity Framework was designed with customisation in mind [203] through an extension mechanism referred to as "Profile". As mentioned, the Framework Profile, also known as Profile,

is one of the main components and strengths of the NIST CSF. NIST defined Profile as "the alignment of the Functions, Categories, and Subcategories with the business requirements, risk tolerance, and resources of the organisation" [195].

Organisations can use Profiles to establish a roadmap for minimising cybersecurity risk that aligns with its objectives, needs, and resources while agreeing with its risk management priorities. These Profiles can be used to describe the current state of various cybersecurity activities as well as the desired target state of those activities. The Current Profile represents the existing/current cybersecurity outcomes, whereas the Target Profile identifies the outcomes that must be accomplished in order to meet the desired cybersecurity risk management objectives [195]. Thus, organisations can detect vulnerabilities in their cybersecurity posture and identify ways for improvement.

Several Profiles have already been released to assist organisations in improving their cybersecurity programs by comparing their current state to these. These profiles represent Target Profiles in specific environments, domains, or events. Some examples of Profiles are:

- the Cybersecurity Framework Manufacturing Profile, which "defines specific cybersecurity activities and outcomes for the protection of the manufacturing system, its components, facility, and environment" [261];
- the Ransomware Risk Management: A Cybersecurity Framework Profile intends to "help organisations to identify and prioritise opportunities for improving their security and resilience against ransomware attacks" [262] as well as to assess their own readiness for this kind of event;
- the Cybersecurity Framework Smart Grid Profile aims to "apply risk management strategies from the NIST CSF to the smart grid" and to "provide cybersecurity risk management guidance to power system owners/operators by prioritising cybersecurity activities based on their effectiveness in helping power system owners/operators achieve common high-level business objectives for the smart grid" [263].

Profiles enable organisations to tailor the NIST CSF activities to specific threats, requirements, and industries. We propose a customisation of the NIST Cybersecurity Framework through the development of a Profile for the management of data manipulation risks. This Profile aims to provide guidance to organisations to prevent and mitigate data manipulation events by aligning their data integrity policy with the elements of the NIST Cybersecurity Framework.

5.3 Conclusion

Cybersecurity frameworks are often generic, unopinionated, and only provide high-level guidance for mitigating cyber risk, leaving it to their users to customise and implement them. In this chapter, we

present a review of a large set of the most complete cybersecurity frameworks using an iterative selection process, with a focus on identifying the most customisable options to manage data integrity threats. This selection process has resulted in a final set of four cybersecurity frameworks. While none of them specifically focus on data integrity threats, we identified the NIST Cybersecurity Framework, a framework designed with customisation, as the best candidate because 1) it offers easy customisation through a native mechanism referred to as “Profile”, and 2) it offers some level of opinion, reducing the overhead to implementation. The customisation mechanism of the NIST CSF allowed us to extend it through the development of a Profile for data manipulation.

CHAPTER 6: CYBERSECURITY FRAMEWORK PROFILE FOR DATA MANIPULATION RISK MANAGEMENT

We identified the NIST Cybersecurity Framework, a framework designed with customization, as the best candidate to support our needs for addressing data manipulation risk. This chapter presents a data manipulation Profile, based on the NIST CSF, that aims to provide guidance and tools to prepare and respond to data tampering incidents. Section 6.1 describes the risks related to data manipulation events. Section 6.2 proposes a Cybersecurity Framework Profile for data manipulation risk management while Section 6.3 presents implementation guidelines.

6.1 The Data Manipulation Challenge

As a result of their digital transformation organisations now have access to and manipulate a considerable amount of digital data. That data has become key to many organisational processes, including decision-making, inviting organisations to protect their data against all kinds of cyber threats.

One of the biggest cybersecurity challenges for an organisation is to continuously protect the integrity of its data. Data integrity refers to the authenticity and trustworthiness of data throughout its life cycle, guaranteeing that the data was/is only modified by authorised and trusted parties. Cyber-attacks are becoming more complex and sophisticated. Cybercriminals have proved their ability to tamper with data and hence compromise the integrity of this data. Data manipulation is a type of cyber-attack that aims to compromise the integrity of data. It is an attack during which a malicious actor performs unauthorised modification to critical data with the intention to manipulate decisions made off that data. These attacks can be just as detrimental to organisations as attacks aiming at stealing data.

A successful data manipulation attack can have serious consequences for organisations. These attacks can be simple, for example changing the value of one or two data at times, yet extremely damaging. This type of attack is difficult to detect and can cast serious doubt on the trust and confidence of an organisation. When sensitive data is compromised, it can lead to corrupted decisions and processes based on that data, harming the trust and confidence at all levels, from employees to partners and customers. Likewise, such integrity-threatening attacks have the potential to compromise entire supply chains, severely harming an organisation's reputation, finances, and more generally, its business. Furthermore, organisations generally take longer to recover from data manipulation attacks since determining the extent and impact of a data manipulation attack, as well as how the data was altered, can be difficult.

Organisations can implement measures to help prevent and counter data manipulation attacks. These measures assist in i) identifying and protecting critical systems, data, and devices; ii) monitoring data and detecting suspicious activity as early as possible, and iii) preparing responses to containing and tracking data manipulation events. The practices outlined in our data manipulation Profile provide an exhaustive strategy to mitigate data manipulation events.

6.2 Data manipulation Profile

The proposed data manipulation Profile aligns security objectives from the NIST Cybersecurity Framework [195] to security measures and practices that support the mitigation, preparation, and response to data manipulation events. This Profile can be used to help organisations manage the risk of data manipulation events or identify opportunities for improving their cybersecurity program to mitigate these risks. This Profile can also be used to assess the readiness of an organisation to mitigate data integrity threats and manage the potential impact of such events.

The Data Manipulation Profile aligns organisations' data integrity prevention and mitigation requirements, objectives, and resources with the elements of the NIST Cybersecurity Framework. The objective of the Profile is to help organisations identify and prioritise opportunities for improving their data integrity policy. This document can be used as a guide for organisations to assess the state of their own readiness.

The Data Manipulation Profile is defined in Table 6.1. The first two columns of the table list the relevant Categories and Subcategories from the NIST CSF. The third column illustrates how each of the Subcategories supports data tampering prevention, detection, and response.

For each Subcategory, the NIST CSF provides additional Informative References. Informative References are specific sections of common critical infrastructure standards, guidelines, and practices that demonstrate a method to achieve the outcomes associated with each subcategory. The Informative References in the Cybersecurity Framework are intended to be informative rather than exhaustive.

The four NIST Cybersecurity Framework Functions used in this Profile to organise the Categories are defined in [195] as:

- **Identify –** Develop an organisational understanding to manage cybersecurity risk to systems, people, assets, data, and capabilities. The activities in the Identify Function are foundational for the effective use of the Framework. Understanding the business context, the resources that support critical functions, and the related cybersecurity risks enables an organisation to focus and prioritise its efforts, consistent with its risk management strategy and business needs.

- **Protect –** Develop and implement appropriate safeguards to ensure the delivery of critical services. The Protect Function supports the ability to limit or contain the impact of a potential cybersecurity event.
- **Detect –** Develop and implement appropriate activities to identify the occurrence of a cybersecurity event. The Detect Function enables the timely discovery of cybersecurity events.
- **Respond –** Develop and implement appropriate activities to take action regarding a detected cybersecurity incident. The Respond Function supports the ability to contain the impact of a potential cybersecurity incident

Due to the criticality and consequences of data manipulation events, it is primordial to identify and contain the tampering as quickly as possible. Therefore, our work focuses on the mean time to identify (MTTI) and mean time to contain (MTTC) this type of events, rather than their recovery. Our Profile focuses on preventing, detecting and responding to data manipulations events. Our Profile only consists of four out of the five Functions of the NIST CSF (e.g., Identify, Protect, Detect and Respond) because these Functions assist in identifying and prioritize critical assets, timely detecting cybersecurity events and containing data manipulation events.

Table 6.1 – Data Manipulation Profile

Category	Subcategory	Data manipulation application
IDENTIFY		
Asset Management (ID.AM): The data, personnel, devices, systems, and facilities that enable the organisation to achieve business purposes are identified and managed consistent with their relative importance to organisational objectives and the organisation's risk strategy.	**ID.AM-1:** Physical devices and systems within the organisation are inventoried **ISO/IEC 27001:2013** A.8.1.1, A.8.1.2 **NIST SP 800-53 Rev. 4** CM-8, PM-5	Asset inventory is an important component of a successful cybersecurity program. Attackers can easily add a new device to your network or take advantage of one of your system's vulnerabilities to tamper your data. A real-time inventory of your assets can help with: i) identifying the physical devices an organisation owns, the potential security risks they represent, and unknown devices, and ii) managing your systems to track known vulnerabilities. Physical device inventories may track information such as device identifier, type, operating system, device owner, physical location, most recent patch applied, and dates of last antivirus definition updates and critical systems updates date. Systems inventories may track information such as system name, identifier, system owner, and system operational status. This information helps to schedule updates and vulnerability scans to reduce the risk of vulnerabilities.
	ID.AM-2: Software platforms and applications within the organisation are inventoried **ISO/IEC 27001:2013** A.8.1.1, A.8.1.2, A.12.5.1 **NIST SP 800-53 Rev. 4** CM-8, PM-5	Asset management applies to software and applications too: i) attackers can access your organisation network or tamper your data via a vulnerable software or application ii) a vulnerable software/application can unintentionally tamper your data. Software and application inventories may document software/application name, version, and owners, devices where it is currently installed, systems associated with the software,

Category	Subcategory	Data manipulation application
		last update date, and current known vulnerabilities. This information helps scheduling security updates and removing dangerous applications and/or software.
	ID.AM-3: Organisational communication and data flows are mapped **ISO/IEC 27001:2013** A.13.2.1, A.13.2.2 **NIST SP 800-53 Rev. 4** AC-4, CA-3, CA-9, PL-8	The growing need for data sharing across (and between) organisations and their different systems increases exposure to cyber threats, especially data manipulation attacks. Data sharing enables corrupted data to travel across systems and organisations. The data flows help understanding, tracking, and containing the propagation of corrupted data within the systems.
	ID.AM-4: External information systems are catalogued **ISO/IEC 27001:2013** A.11.2.6 **NIST SP 800-53 Rev. 4** AC-20, SA-9	Data is not always stored internally and often resides on external information systems (e.g., from a partner or cloud provider). Even if these systems are external to your organisation, they need to be cataloged because they still represent potential attack points. Attackers can use external systems as an opportunity to propagate tamper data within your system with disastrous consequences for your organisation. In addition to data flows, this inventory helps to identify the connections between internal and external systems, to document the source and destination of data exchanges, and to facilitate tracking of corrupted data.
	ID.AM-6: Cybersecurity roles and responsibilities for the entire workforce and third-party stakeholders (e.g., suppliers, customers, partners) are established	Data integrity can be compromised by various factors. One of the most basic and common is human error, for example by not following the organisation's security guidelines. It is therefore important to educate everyone in the organisation on security policies and their role and responsibilities to help with

Category	Subcategory	Data manipulation application
	ISO/IEC 27001:2013 A.6.1.1 **NIST SP 800-53 Rev. 4** CP-2, PS-7, PM-11	prevention, reporting, and containment of data manipulation in a timely fashion.
Risk Assessment (ID.RA): The organisation understands the cybersecurity risk to organisational operations (including mission, functions, image, or reputation), organisational assets, and individuals.	**ID.RA-1:** Asset vulnerabilities are identified and documented **ISO/IEC 27001:2013** A.12.6.1, A.18.2.3 **NIST SP 800-53 Rev. 4** CA-2, CA-7, CA-8, RA- 3, RA-5, SA-5, SA-11, SI-2, SI-4, SI-5	Attackers can take advantage of an asset vulnerability to conduct an attack. Identifying and documenting the vulnerabilities of the organisation's assets helps to better protect your assets and mitigate those vulnerabilities to reduce risks. It also helps to develop plans to assess and respond to future data manipulation events.
	ID.RA-3: Threats, both internal and external, are identified and documented **ISO/IEC 27001:2013** Clause 6.1.2 **NIST SP 800-53 Rev. 4** RA-3, SI-5, PM-12, PM- 16	Data integrity is subject to various potential threats. Threat assessment is important to (adapt and) consolidate security strategies and reduce the risk of data manipulation. Identifying and documenting threats helps the organisation understand and become familiar with the techniques used to target the organisation and therefore develop plans to prevent unauthorised access and data breaches, and develop awareness programs to train all the stakeholders against organisation-specific threats.
	ID.RA-5: Threats, vulnerabilities, likelihoods, and impacts are used to determine risk **ISO/IEC 27001:2013** A.12.6.1 **NIST SP 800-53 Rev. 4** RA-2, RA-3, PM-16	Data integrity attacks can be dangerous for organisations. Data integrity is exposed to many potential risks that you should identify and understand. Determining integrity risk relies on assessing security threats, identifying systems vulnerabilities, and determining the impacts of data manipulation attacks. Understanding data integrity risks should help your organisation prioritise its protection efforts. Identifying and

Category	Subcategory	Data manipulation application
		understanding risks also helps develop plans to manage them, minimise their impacts, and create awareness of those risks within the organisation.
Supply Chain Risk Management (ID.SC): The organisation's priorities, constraints, risk tolerances, and assumptions are established and used to support risk decisions associated with managing supply chain risk. The organisation has established and implemented the processes to identify, assess and manage supply chain risks	**ID.SC-5:** Response and recovery planning and testing are conducted with suppliers and third-party providers **ISO/IEC 27001:2013** A.17.1.3 **NIST SP 800-53 Rev. 4** CP-2, CP-4, IR-3, IR-4, IR-6, IR-8, IR-9	Most organisations work with suppliers and third-party providers who are exposed to consuming or propagating tampered data. Your organisation can potentially share compromised data with them before the attack is detected. It is critical to include suppliers and third-party providers in the development of response and recovery plans because of a great chance that they will be impacted as well.
PROTECT		
Identity Management, Authentication and Access Control (PR.AC): Access to physical and logical assets and associated facilities is limited to authorised users, processes, and devices, and is managed consistent with the assessed risk of unauthorised access to authorised activities and transactions.	**PR.AC-1:** Identities and credentials are issued, managed, verified, revoked, and audited for authorised devices, users and processes **ISO/IEC 27001:2013** A.9.2.1, A.9.2.2, A.9.2.3, A.9.2.4, A.9.2.6, A.9.3.1, A.9.4.2, A.9.4.3 **NIST SP 800-53 Rev. 4** AC-1, AC-2, IA-1, IA-2, IA-3, IA-4, IA-5, IA-6, IA-7, IA-8, IA-9, IA-10, IA-11	Most data manipulation attacks start with credential compromise and result from unauthorised access (or exposure) to files or systems. A strong authentication and credential management protocols is an essential step to mitigate those risks.

Category	Subcategory	Data manipulation application
	PR.AC-4: Access permissions and authorisations are managed, incorporating the principles of least privilege and separation of duties **ISO/IEC 27001:2013** A.6.1.2, A.9.1.2, A.9.2.3, A.9.4.1, A.9.4.4, A.9.4.5 **NIST SP 800-53 Rev. 4** AC-1, AC-2, AC-3, AC- 5, AC-6, AC-14, AC-16, AC-24	Data manipulation attacks occur when individuals (within or outside an organisation) get access to systems or data even though they do not have proper authorisation. To prevent unauthorised change to the data, users should be granted access to data based on their roles, responsibilities, and work they perform.
	PR.AC-5: Network integrity is protected (e.g., network segregation, network segmentation) **ISO/IEC 27001:2013** A.13.1.1, A.13.1.3, A.13.2.1, A.14.1.2, A.14.1.3 **NIST SP 800-53 Rev. 4** AC-4, AC-10, SC-7	Network security is essential when working over the internet because it helps protect your data against integrity attacks. If your network integrity is compromised, then attackers have access to your data, and it is most likely compromised as well. Network segmentation and segregation can minimise the impact of network intrusion, unauthorised access to data, and reduce the propagation of tampered data.
	PR.AC-6: Identities are proofed and bound to credentials and asserted in interactions **ISO/IEC 27001:2013**, A.7.1.1, A.9.2.1 **NIST SP 800-53 Rev. 4** AC-1, AC-2, AC-3, AC- 16, AC-19, AC-24, IA-1, IA-2, IA-4, IA-5, IA-8, PE-2, PS-3	Most data manipulation attacks can be linked to compromised credentials, therefore identities should be proofed and bound to credentials to prevent unauthorised access.

Category	Subcategory	Data manipulation application
Awareness and Training (PR.AT): The organisation's personnel and partners are provided cybersecurity awareness education and are trained to perform their cybersecurity- related duties and responsibilities consistent with related policies, procedures, and agreements.	**PR.AT-1:** All users are informed and trained **ISO/IEC 27001:2013** A.7.2.2, A.12.2.1 **NIST SP 800-53 Rev. 4** AT-2, PM-13	Data integrity can be compromised through human error, malicious (e.g., unauthorised access) or unintentional (e.g., enter incorrect data, duplicate data), and those breaches can be difficult to identify. Organisations should teach all users how to identify and correctly respond to cyber threats as well as provide training about the importance of data integrity. Awareness training should help users understand how they can contribute to protecting the quality of the organisation's data.
	PR.AT-2: Privileged users understand their roles and responsibilities **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2 **NIST SP 800-53 Rev. 4** AT-3, PM-13	Privileged users have unlimited access and permissions to critical systems and data within the organisations making them targets for cybersecurity criminals or insider threats. Privileged users need to understand their roles and responsibilities with regards to data integrity.
	PR.AT-4: Senior executives understand their roles and responsibilities **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2 **NIST SP 800-53 Rev. 4** AT-3, PM-13	The organisation provides training to the personnel that is adapted to their roles and responsibilities. Senior executives need to be aware of their role and responsibilities regarding security in order to help their subordinates protect data integrity and mitigate risks.
	PR.AT-5: Physical and cybersecurity personnel understand their roles and responsibilities **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2 **NIST SP 800-53 Rev. 4** AT-3, IR-2,	Cybersecurity personnel are the first line of defys against cyber-attacks and are responsible for preventing, monitoring, and responding to data breaches. Therefore, they need to understand their roles in order to protect organisation systems and data, identify threats and help other users against cyber menaces.

Category	Subcategory	Data manipulation application
	PM-13	
Data Security (PR.DS): Information and records (data) are managed consistent with the organisation's risk strategy to protect the confidentiality, integrity, and availability of information.	**PR.DS-1:** Data-at-rest is protected **ISO/IEC 27001:2013** A.8.2.3 **NIST SP 800-53 Rev. 4** MP-8, SC-12, SC-28	Data manipulation attacks can happen to data-at-rest as well as to data-in-transit. Data needs to be protected at any time (before, during, and after exchange) to avoid compromising data integrity. A strategy to detect and track modifications of data-at-rest should be defined and implemented.
	PR.DS-2: Data-in-transit is protected **ISO/IEC 27001:2013** A.8.2.3, A.13.1.1, A.13.2.1, A.13.2.3, A.14.1.2, A.14.1.3 **NIST SP 800-53 Rev. 4** SC-8, SC-11, SC-12	Data manipulation attacks can happen to data-at-rest as well as to data-in-transit. Data needs to be protected at any time (before, during, and after exchange) to avoid compromising data integrity. A strategy to validate data integrity after transit (i.e., confirm that data was not tampered during the exchange) should be defined and implemented.
	PR.DS-6: Integrity checking mechanisms are used to verify software, firmware, and information integrity **ISO/IEC 27001:2013** A.12.2.1, A.12.5.1, A.14.1.2, A.14.1.3, A.14.2.4 **NIST SP 800-53 Rev. 4** SC-16, SI-7	Data integrity checking mechanisms should be used to prevent using or exchanging/propagating tampered data. These mechanisms should support data-at-rest and data-in-transit.
DETECT		
Anomalies and Events (DE.AE): Anomalous activity is detected, and the potential impact of events is	**DE.AE-2:** Detected events are analysed to understand attack targets and methods	Analysing detected data manipulation events allows the organisation to learn from the targets chosen and the methods used for data tampering in order to improve and prioritise data

Category	Subcategory	Data manipulation application
understood.	**ISO/IEC 27001:2013** A.12.4.1, A.16.1.1, A.16.1.4 **NIST SP 800-53 Rev. 4** AU-6, CA-7, IR-4, SI-4	protection as well as improve response and recovery plans.
	DE.AE-4: Impact of events is determined **ISO/IEC 27001:2013** A.16.1.4 **NIST SP 800-53 Rev. 4** CP-2, IR-4, RA-3, SI-4	Evaluating the impact of data manipulation events can help organisations to identify their most sensitive data, understand the consequences it can have on the organisation if this data is compromised, and how this tampered data can propagate within and between organisations.
Security Continuous Monitoring (DE.CM): The information system and assets are monitored to identify cybersecurity events and verify the effectiveness of protective measures.	**DE.CM-1:** The network is monitored to detect potential cybersecurity events **NIST SP 800-53 Rev. 4** AC-2, AU-12, CA-7, CM- 3, SC-5, SC-7, SI-4	Data integrity attacks can be difficult to detect. In some cases, data manipulation attacks can go undetected for years before data is identified as tampered with. That is why it is important to track any suspicious activities on the network to prevent data tampering. Continuous network monitoring might detect intrusions before any data tampering can be done to compromise the integrity of the information and before causing major damage to the organisation.
	DE.CM-3: Personnel activity is monitored to detect potential cybersecurity events **ISO/IEC 27001:2013** A.12.4.1, A.12.4.3 **NIST SP 800-53 Rev. 4** AC-2, AU-12, AU-13, CA-7, CM-10, CM-11	Data integrity attacks often result from human error, accidentally or intentionally. It is critical to monitor any attempts (read, modify or delete) to access sensitive data, and to assess (where, when, how) the usual activities of the users accessing data. Continuous monitoring of personnel activity might detect insider threats, unsafe personnel practices (sudden increase of activity, activity outside business hours), or compromised

Category	Subcategory	Data manipulation application
		credentials that can lead to data manipulation attacks.
	DE.CM-7: Monitoring for unauthorised personnel, connections, devices, and software is performed **ISO/IEC 27001:2013** A.12.4.1, A.14.2.7, A.15.2.1 **NIST SP 800-53 Rev. 4** AU-12, CA-7, CM-3, CM-8, PE-3, PE-6, PE-20, SI-4	Continuous monitoring limits unauthorised access to your sensitive data. Continuous monitoring can detect data manipulation (and data integrity) attacks before they happen, by identifying unauthorised devices, connections, and/or personnel. Monitoring help ensure that the data stays secure and keeps its integrity.
	DE.CM-8: Vulnerability scans are performed **ISO/IEC 27001:2013** A.12.6.1 **NIST SP 800-53 Rev. 4** RA-5	Data integrity attacks can start from a vulnerability in your system or network that you don't know about or that was not remediated. New vulnerabilities are being identified every day, and you should always be up-to-date with all the vulnerabilities in your organisation's systems. Regular scans allow you to detect these new vulnerabilities and help prioritise the mitigation of those vulnerabilities before they can be used to launch data manipulation attacks.
RESPOND		
Communications (RS.CO): Response activities are coordinated with internal and external stakeholders (e.g. external support from law enforcement agencies).	**RS.CO-2:** Incidents are reported consistent with established criteria **ISO/IEC 27001:2013** A.6.1.3, A.16.1.2 **NIST SP 800-53 Rev. 4** AU-6, IR-6, IR-8	Users that access the data regularly can detect data tampering. This can be the result of an error or an attack. Personnel should be trained on the different methods to report incidents. Organisations should define pre-established criteria and forms for reporting incidents. Users' reports should be used to improve the response plan and provide metrics on the incident response capabilities.

Category	Subcategory	Data manipulation application
	RS.CO-3: Information is shared consistent with response plans **ISO/IEC 27001:2013** A.16.1.2, Clause 7.4, Clause 16.1.2 **NIST SP 800-53 Rev. 4** CA-2, CA-7, CP-2, IR-4, IR-8, PE-6, RA-5, SI-4	When an integrity attack is detected, it is key to share information about the event with partners to help prevent the spread and use of corrupted data to other systems or organisations. Consistent sharing also improves the coordination of response teams and the prevention of data manipulation attacks.
	RS.CO-5: Voluntary information sharing occurs with external stakeholders to achieve broader cybersecurity situational awareness **ISO/IEC 27001:2013** A.6.1.4 **NIST SP 800-53 Rev. 4** SI-5, PM-15	The massive use of digital networks has accentuated the diversity of cyber-attacks. To better protect your organisation's sensitive data, information sharing with external stakeholders can help to get a better understanding of cyber-menaces. As a result of this information sharing, organisations can better adapt their cybersecurity programs to new threats and mitigate the risks of data manipulation attacks.
Mitigation (RS.MI): Activities are performed to prevent expansion of an event, mitigate its effects, and resolve the incident.	**RS.MI-3:** Newly identified vulnerabilities are mitigated or documented as accepted risks **ISO/IEC 27001:2013** A.12.6.1 **NIST SP 800-53 Rev. 4** CA-7, RA-3, RA-5	Attackers will use every vulnerability of your system to access your sensitive data and tamper it. When new vulnerabilities are identified, and if no immediate mitigation is possible, data consumers should be made aware of the risk(s) associated with consuming the data.

6.3 Implementation guidelines

In complement of the data manipulation Profile, we propose implementation recommendations in order to provide a comprehensive solution for data traceability. Those recommendations are composed of a collection of users stories for data traceability, a set of security metrics and a proposition for an information flow awareness training.

6.3.1 Data traceability

The digital transformation enabled the generation of large volumes of digital data as well as the development of systems and software to manage it. The amount of data that the manufacturing organisations produce and consume has clearly risen in the last decade and it will most likely continue to grow in the future, with the developing reliance on cyber-physical systems and IoT sensors. Besides, organisations rely on new technologies to exchange and integrate data within or across business units or organisations. This digital transformation has improved collaboration between all the stakeholders, systems, and machines. For example, instructions and designs can be shared on the cloud or distributed instantly to thousands of people at once, or machines can be monitored in real-time using wireless sensors [184]. The development of artificial intelligence (AI) and machine learning (ML) for data analytics lead to more educated decisions and better scheduling for optimised productivity, among others. These improvements are due to the availability of data in a computer processable format, as well as the quality and integrity of data that are key drivers of data-driven decisions. Trusted data is critical for making and supporting educated decisions throughout the product lifecycle, from design to inspection.

However, as mentioned, this digital transformation makes organisations more exposed to digital threats. Smart manufacturing relies on data to support collaboration and decision-making, the cyber threats can have numerous and significant consequences. Data tampering has been confirmed to be a serious threat to manufacturers and the products they build [184].

Effective collaboration requires unambiguous information flow across the systems that produce and consume the data. Data must be accessible and with enough context to be interpretable, across business units within an organisation or across business units from different organisations. Different systems must be able to understand the information. Data flows provide a complete view of the data exchange between systems and processes within an organisation or supply chain.

The complexity of product and organisation is proportional to the need and use of collaboration and hence to the number of data exchanges. Data exchanges are key enablers for propagating tampered data and spreading corrupted information in the different systems of an organisation. The advances in information

technologies facilitate this propagation, making it happen at a faster rate and higher reach. For example, compromised files can be instantly shared with suppliers on different continents or corrupted sensors can report incorrect data in real-time. Inaccurate decisions or defective products can result from processing this altered data. It is critical to determine the source of the tampering as soon as possible in order to rapidly identify the systems and decisions at risk and thus, effectively contain the threat.

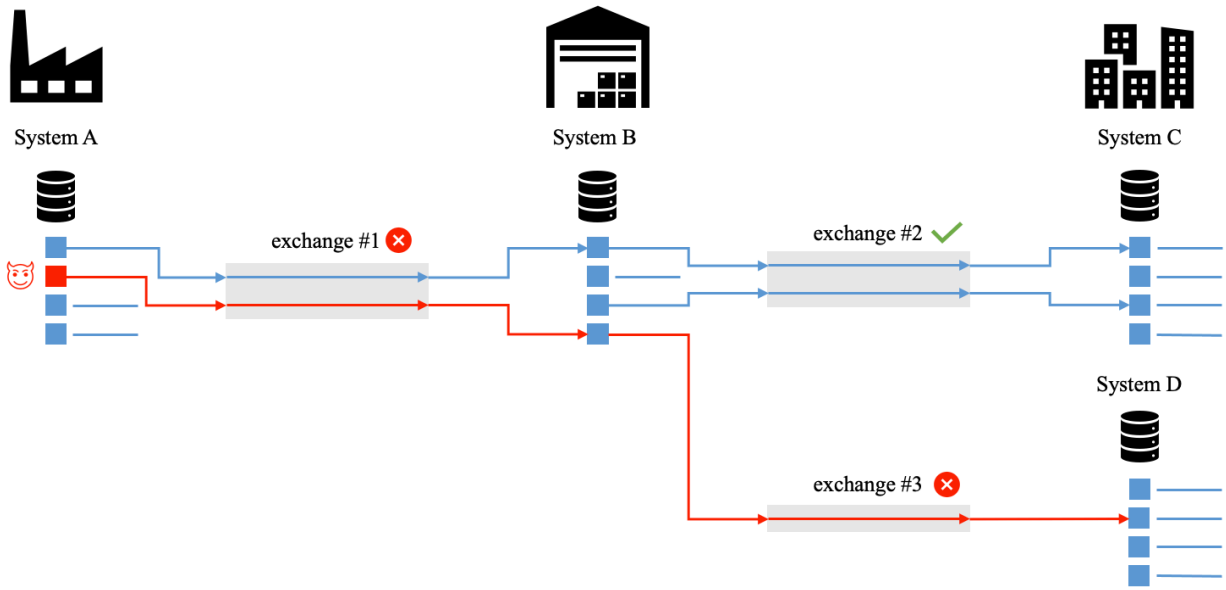

Figure 6.1 – Graphical representation of the propagation of tampered data and the role of data mapping to identify that propagation. [184]

Solutions like digital signatures can be used to identify the organisation(s) potentially at risk, but for information systems, it is necessary to identify the system(s) that are consuming and potentially propagating tampered data in order to isolate them and thus correctly contain the threat. During a data manipulation attack, not all data fields are automatically compromised, only the fields of the system(s) consuming or exchanging the tampered data are concerned. To control the propagation of the threat and avoid a complete shutdown of all the systems, it is essential to precisely identify the infected system(s). To do so, a detailed view of what is consumed and exchanged by all systems, at the data field level, is necessary, otherwise, it is impossible to determine which system is compromised. For example, in the scenario depicted in Figure 6.1, the systems of three organisations are exchanging data, and system A is corrupted. A precise definition and view of the exchanges, represented by the arrows inside the blue exchange boxes, enables to easily identify that between the systems C and D, the only system compromised is system D. In Figure 6.1, the blue boxes represent the data mappings that are required to correctly connect the data fields of the different systems. Data mappings are critical for a digital enterprise as it enables heterogeneous systems to

communicate. A mapping specifies the equivalence and associated transformations required between data representations (i.e., data schemes).

Digital collaboration is optimal when all of the systems involved use structured or semi-structured data with unambiguous definitions and data schemes because the lack of ambiguity in the definition provides for a clear understanding of the data and a successful mapping between the definitions. The complexity of a mapping definition greatly depends on a few factors, such as quantity, structural distance, business rules alignment and reference data libraries, related to the source and target schemes involved in the exchange.

Data mapping, which specifies how systems are connected and represents the data flows between them, is a critical component to detect and contain the propagation of tampered data. These mapping specifications should be examined, following the detection of a threat, to identify the systems that are contaminated and potentially contaminating others, as well as the information, decisions, and products at risk.

Through data mapping, organisations can model and trace the data flows within their supply chain in order to determine their exposure to data manipulation threats. Mapping and monitoring data exchange would help organisations identify risks and vulnerabilities and thus better protect their assets against threats and contain the propagation of corrupted data in case of attacks.

6.3.2 User stories

User stories are short and simple descriptions of a feature from the perspective of a user or customer of the system, stated in an informal and natural language. They are one of the key components of Agile software development and are designed to capture user requirements. They represent high-level activities users can do in the future product. We defined a non-exhaustive collection of user stories representing features that should be implemented to achieve effective data traceability. These user stories align with the Categories and Subcategories of our data manipulation Profile and some examples of these user stories are detailed in the Table below.

These user stories define informal and general descriptions of a feature for our data traceability solution written from the perspective of different users such as:

- **Chief information security officer (CISO)** is responsible for defining and implementing the IT security strategy, as well as protecting the organisation against security threats.
- **Incident response team** is responsible for incident management (i.e., investigating and responding to security incidents) as well as developing security strategies for threat prevention.

- **Information system administrator** is responsible for managing and updating hardware and software assets, as well as implementing and managing access controls.
- **IT asset manager** is responsible for managing and protecting the organisation's IT assets throughout their life cycle.
- **IT contingency planning coordinator** is responsible for developing and maintaining the contingency plan.
- **IT security program manager** is responsible for developing or applying an appropriate methodology to help identify, evaluate, and minimise IT security risks to the organisation as well as performing system risk analyses.
- **IT system security officer** is responsible for ensuring the security of an information system.
- **Network security team** is responsible for managing the network security by defining and implementing strategies to protect the network, as well as monitoring network vulnerabilities.
- **Senior agency information security officer (SAISO)** is responsible for the organisation's IT/cybersecurity awareness and training program.
- **Security Operations Center (SOC)** is responsible for the organisation's overall cybersecurity through incident detection, prevention and response, risk management, and monitoring, among other things. SOCs have a much broader scope of responsibility than incident response teams that are more specialised.
- **Training developer** is responsible for developing role-based cybersecurity training material and courses.

Table 6.2 – User stories for the data manipulation Profile

Category	Subcategory	User stories
IDENTIFY		
Asset Management (ID.AM)	**ID.AM-1:** Physical devices and systems within the organisation are inventoried	• As the IT asset manager, I want to have a centralised repository for the inventory of my organisation's assets (systems, hardware, software), so I can rapidly identify assets that have been compromised, breached, or are otherwise in need of mitigation actions. • As the IT asset manager, I want to establish an inventory of all systems, devices, and software within the organisation, so I can monitor the organisation's assets, and anticipate events that can affect employees' work (e.g., software updates that can change how it works) or the organisation's security (e.g., lost devices, external systems not fully patched). • As the IT asset manager, I want to review and update the assets inventory so I can ensure that this inventory is up-to-date, accurate, and complete at any time. • As the IT asset manager, I want to assign devices to systems as well as software and applications to devices and systems, so I can trace dependencies. • As the IT asset manager, I want the inventory information to include the individuals responsible and accountable for administering each asset, so I can easily identify and contact those individuals if some action is required (e.g., if the asset is identified as the source of a breach or needs to be replaced).
	ID.AM-2: Software platforms and applications within the organisation are inventoried	
	ID.AM-3: Organisational communication and data	• As the information system administrator, I want to map the data flows (both internal and external) of the system(s) I manage, so I can ensure the traceability of data produced and consumed by these systems.

Category	Subcategory	User stories
	flows are mapped	• As the information system administrator, I want to update any data flows of the system(s) I manage, so I keep the data mapping up-to-date and the data traceability accurate. • As the *incident response team*, I want to access the data mapping, so I can trace and analyse the data flows in case of data manipulation attacks. • As a Senior Information Security Officer, I want to access the mapping, so I can use them as material for the Information Flow Awareness Training (see in section 6.3.4).
	ID.AM-4: External information systems are catalogued	• As the network security teams (or the network manager), we want to document all external systems use within the organisation, so I can monitor and control the external systems that have access to the organisation's systems and data. • As the network security team (or the network manager), we want to review and update the inventory of the external systems, so I can have an up-to-date, accurate and complete inventory to track the authorised systems that can access your organisation in real-time. • As the network security teams (or the network manager), we want the inventory information to include the name(s) of the organisation and/or individual that own the external system, so I can control and contact those entities/individuals if some action is required.
	ID.AM-6: Cybersecurity roles and responsibilities for the entire workforce and third-party stakeholders (e.g., suppliers, customers, partners) are established	• As the Senior Information Security Officer, I want to establish and maintain security awareness and training programs, so I can ensure that each employee has information to understand their role and responsibilities. • As a user, I want to access training materials, so I can maintain awareness about my role and responsibilities within the organisation

Category	Subcategory	User stories
Risk Assessment (ID.RA)	**ID.RA-1:** Asset vulnerabilities are identified and documented	• As the Security Operations Center, we want to monitor and scan the systems, so we can identify and report vulnerabilities that can affect the organisation. • As the IT Security Program Manager, I want to implement vulnerability management tools, so I can identify, mitigate and minimise the opportunities for attackers. • As the Security Operations Center, we want to record the vulnerabilities into a centralised document, so we can easily update and review existing vulnerabilities.
	ID.RA-3: Threats, both internal and external, are identified and documented	• As the IT Security Program Manager, I want to conduct risk assessments, so I can identify threats and document the likelihood and impact they can have on the organisation. • As members of the Security Operations Center or the incident response teams, we want to establish and maintain a cyber threat hunting, so we can detect, track and mitigate threats that can escape existing controls. • As the Security Operations Center, we want to record the threats into a centralised document, so we can easily update and review existing threats.
	ID.RA-5: Threats, vulnerabilities, likelihoods, and impacts are used to determine risk	• As the IT Security Program Manager, I want to conduct risk assessments, so I can identify, estimate and prioritise risks within the organisation. • As the IT Security Program Manager, I want to develop risk assessment reports, so they can be used to support decision-making.
Supply Chain Risk Management (ID.SC)	**ID.SC-5:** Response and recovery planning and testing are conducted with suppliers and third-party providers	• As the IT Contingency Planning Coordinator, I want to coordinate contingency planning with external parties, so I can develop a comprehensive and timely plan that reflects the overall response and recovery needs of the organisation. • As the IT Contingency Planning Coordinator, I want to distribute the organisation's contingency and incident response plans to external parties,

Category	Subcategory	User stories
		so I can ensure that suppliers and third-party providers have access to the guidelines defined by the organisation to process after a security event. • As the IT Contingency Planning Coordinator, I want to communicate the organisation's contingency and incident response plans changes to external parties, so I can ensure that suppliers and third-party providers have an up-to-date and accurate version of the organisation's plans.
PROTECT		
Identity Management, Authentication and Access Control (PR.AC)	**PR.AC-1:** Identities and credentials are issued, managed, verified, revoked, and audited for authorised devices, users and processes	• As the Information System Administrator, I want to assign unique identifiers to users, so I can track and monitor users. • As the IT asset manager, I want to assign unique identifiers to devices and processes, so I can track and monitor devices and processes. • As an IT System Security Officer, I want to authenticate devices before connection, so I can track and monitor the devices accessing the system(s) I am responsible for. • As an Information System administrator, I want to approve users, so I can grant them access to the resource I am responsible for.
	PR.AC-4: Access permissions and authorisations are managed, incorporating the principles of least privilege and separation of duties	• As an Information System administrator, I want to define an access control policy, so I can limit system access to authorised users and devices. • As an Information System administrator, I want to manage access permissions in the system(s) I am responsible for, so I can authorise and/or limit access to users according to their assigned organisational tasks. • As an IT System Security Officer, I want to document authorised users for the system(s) I am responsible for, so I can track and monitor those users. • As an Information System administrator, I want to perform access reviews regularly, so I can control that access is still required and if necessary, remove access within a short timeframe.

Category	Subcategory	User stories
	PR.AC-5: Network integrity is protected (e.g., network segregation, network segmentation)	• As the network security team, we want to define security policies for implementing network segmentation and segregation, so we can protect network integrity. • As the CISO, I want to conduct assessments, so I can ensure that network segments are correctly implemented and operated. • As the Security Operations Center, we want to implement network monitoring tools, so we can monitor and control communication and network activities, and thus prevent breaches.
	PR.AC-6: Identities are proofed and bound to credentials and asserted in interactions	• As an Information System administrator, I want to periodically verify the identity of users accessing the system(s) I am responsible for, so I can minimise the likelihood of a breach. • As data owners, we want to periodically verify that only necessary users are granted access to the sensible data we are responsible for, so we can reduce the risk of a data breach. • As an Information System administrator, I want to periodically review the list of all privileged access to the system(s) I am responsible for, so I can ensure visibility and verification for privileged access accounts in order to reduce the risk of a data breach. • As an Information System administrator, I want to manage access permissions in the system(s) I am responsible for, so I can authorise and/or limit access to users according to their assigned organisational tasks.
Awareness and Training (PR.AT)	**PR.AT-1:** All users are informed and trained	• As a training developer and the Senior Information Security Officer, we want to define role-based training, so we can ensure that each employee is trained and educated according to their assigned roles and responsibilities.

Category	Subcategory	User stories
	PR.AT-2: Privileged users understand their roles and responsibilities	• As a training developer and the Senior Information Security Officer, we want to update the security training, so we can ensure that each employee is up-to-date with the cybersecurity threats and policies. • As a user, I want to access training materials so I can learn and understand the need for security and the actions (e.g., recognising and reporting) to take to maintain security. • As a user, I want to get notified when new training materials are available, so I can review updated training and stay up-to-date with the organisation's security policy. • As the CISO, I want to get reports on the progress and completion of the employees' training, so I can monitor the course completion rates, the level of engagement, and the level of knowledge retained.
	PR.AT-4: Senior executives understand their roles and responsibilities	
	PR.AT-5: Physical and cybersecurity personnel understand their roles and responsibilities	
Data Security (PR.DS)	**PR.DS-1:** Data-at-rest is protected	• As the network security team, we want to implement cryptographic mechanisms, so I can prevent unauthorised access or modifications to both data-at-rest and data-in-transit. • As the CISO, I want to approve the cryptographic mechanisms, so I can ensure both data-at-rest and data-in-transit are encrypted and thus protected. • As the network security teams (or the network manager), we want to implement network security controls like firewalls and network access control, so we can help secure the storage devices and networks against intrusions and other cyber threats.
	PR.DS-2: Data-in-transit is protected	
	PR.DS-6: Integrity checking mechanisms are used to verify software, firmware, and information integrity	• As the IT Security Program Manager, I want to implement integrity verification tools, so I can detect unauthorised changes to software, firmware, and information. • As members of the Security Operations Center, we want to investigate and resolve integrity check alerts, so we can ensure that the integrity of the data processed is still intact.

Category	Subcategory	User stories
		• As the IT Security Program Manager, I want to define actions to take when unauthorised changes to software, firmware, and information, so I can reduce response time and limit propagation.
DETECT		
Anomalies and Events (DE.AE)	**DE.AE-2:** Detected events are analysed to understand attack targets and methods	• As members of the Security Operations Center, we want to review system audit logs, so we can analyse and investigate suspicious and unauthorised systems activities. • As members of the Security Operations Center, we want to implement monitoring tools, so we can identify and observe events occurring within the system. • As members of the Security Operations Center, we want to investigate alerts from monitoring tools, so we can resolve or escalate events to the CISO as appropriate.
	DE.AE-4: Impact of events is determined	• As members of the Security Operations Center, we want to review the critical systems and dependencies catalogued by the IT asset manager, so we can determine the impact of security events. • As members of the Security Operations Center, we want to conduct incident analysis, so we can evaluate the severity of the threat and the impact it may have on the organisation in order to formulate an appropriate response strategy. • As members of the Security Operations Center, we want to investigate incidents, so we can triage and prioritise them based on the sensitivity of the data and assets involved. • As the IT Security Program Manager, I want to conduct a risk assessment, so I can determine the likelihood and impact of events.
Security Continuous	**DE.CM-1:** The network is	• As the Security Operations Center (with the help of the IT Security

Category	Subcategory	User stories
Monitoring (DE.CM)	monitored to detect potential cybersecurity events	Program Manager), we want to develop and implement a continuous monitoring strategy, so we can detect attacks or indicators of potential attacks. • As the Security Operations Center, we want to implement monitoring tools, so we can identify unauthorised use of organisational systems. • As an Information System administrator, I want to control and monitor user-installed software, so I can limit the installation of unauthorised software on the organisation's devices as they bring new vulnerabilities to the systems.
	DE.CM-3: Personnel activity is monitored to detect potential cybersecurity events	• As the Security Operations Center (with the help of the IT Security Program Manager), we want to develop and implement a continuous monitoring strategy, so we can detect suspicious activities. • As the Security Operations Center, we want to implement user activities monitoring and network protection tools, so we can monitor and track end user behaviour on devices, networks, and other organisation's systems. • As the Security Operations Center, we want to monitor and control communication, so we can detect suspicious communications or activities. • As the Security Operations Center, we want to create, preserve, and review system audit logs, so we can monitor, analyse, and investigate unlawful or unauthorised system activity.
	DE.CM-7: Monitoring for unauthorised personnel, connections, devices, and software is performed	• As the Security Operations Center, we want to implement user activities and system monitoring tools, so we can identify suspicious behaviours and activities. • As members of the Security Operations Center, we want to analyse system audit logs, so we can detect unauthorised users, connections, or devices. • As members of the Security Operations Center, we want to investigate alerts of unknown assets connection or unauthorised use of the

Category	Subcategory	User stories
		organisation's resources. • As members of the Security Operations Center, we want to monitor physical access to the organisation, so we can detect unauthorised access to devices or information.
	DE.CM-8: Vulnerability scans are performed	• As the Security Operations Center, we want to regularly monitor and scan for vulnerabilities in systems, so we can identify and report new vulnerabilities. • As the Security Operations Center, we want to store vulnerability scan reports in a centralised storage device, so we can easily access those reports at any time. • As the CISO, I want to access vulnerability scan reports, so I can ensure that the scheduled vulnerability scans are completed.
RESPOND		
Communications (RS.CO)	**RS.CO-2:** Incidents are reported consistent with established criteria	• As the CISO (with the help of the incident response team), I want to develop an incident response plan, so users can report suspected incidents in a consistent fashion. • As a user, I want to report an incident so I can warn the cybersecurity team about unusual activity. • As a user, I want to consult the tickets I submitted, so I can see the progress of my tickets. • As a member of the Security Operations Center, we want to have access to the list of the tickets, so we can do the triage and assign tickets to a specific member(s) of the incident response team. • As the Security Operations Center, we want to track and document incidents, so we can maintain records about each accident reported within the organisation. • As a member of the response team, I want to access any tickets assigned

Category	Subcategory	User stories
		to me, so I can investigate them.
	RS.CO-3: Information is shared consistent with response plans	• As the CISO (with the help of the incident response team), I want to develop an incident response plan to define steps for escalating incidents to the executive response team, so IT teams have guidelines to mitigate and respond quickly and uniformly to threats. • As the incident response team, we want to investigate incidents, so we can determine the appropriate stakeholders to contact and when to communicate information about the incident.
	RS.CO-5: Voluntary information sharing occurs with external stakeholders to achieve broader cybersecurity situational awareness	
Mitigation (RS.MI): Activities are performed to prevent expansion of an event, mitigate its effects, and resolve the incident.	**RS.MI-3:** Newly identified vulnerabilities are mitigated or documented as accepted risks	• As the Security Operations Center, we want to periodically scan systems for vulnerabilities, so we can identify new vulnerabilities. • As the incident response team, we want to develop and implement plans of action, so we can mitigate new vulnerabilities and reduce risks. • As the Security Operations Center, we want to update the vulnerability document as new vulnerabilities are discovered, so IT security teams can always be up-to-date with the vulnerabilities of the systems. • As the CISO, I want to review vulnerability reports, so I can ensure that all vulnerabilities are mitigated in a timely manner.

6.3.3 Security Metrics

By definition, metrics are quantifiable measures used to monitor and evaluate all the different parts of a business. Implementing metrics helps organisations improve their overall performances and align their employees and processes with their business objectives. The goal of security metrics is to determine what needs to be improved in an organisation's security policies, practices, processes, and technology, in order to assess its readiness against cyber-attacks.

The security metrics we are proposing aim to assess organisations' readiness against data integrity and data manipulation attacks and manage their risk exposure to those threats. Our set of security metrics is composed of 12 metrics, which are split into five distinct categories:

- Data vulnerabilities: it focuses on identifying vulnerable data in the system and analysing their impact on the systems and network;
- Access management: it aims to identify, track and manage users' access to systems and data;
- Monitoring and response: it focuses on assessing vulnerability reports and/or detections, and analysing responses to threats;
- Network metrics: it deals with the analysis of the vulnerabilities of the network;
- Security assessment: it focuses on analysing the success of awareness and security training programs

Note: We consider that, in the following security metrics, a class in a system is an information class in a class diagram and that an object or element of a class is an instance of this class.

6.3.3.1 Data vulnerabilities

Connecting Corruption Propagation (CCP)

In an organisation, multiple systems communicate with one another and thus can be reliant on the data exchanged. Data flows enable to represent the systems that are exchanging data and thus the systems that are connected. This metric aims to identify the systems that will impact others the most (i.e., how far potential tampered data may propagate to other systems) and to estimate the extent of damage that could occur if some data were tampered with. This metric represents the number of systems that could be affected by corrupted data in a specific system. Each system has a "propagation lane" which corresponds to all the systems that are likely to receive, directly or indirectly data from this system. It helps increase the security and monitoring of vulnerable systems, which are the systems with the longest "propagation lane" and thus are the most likely to propagate tampered data within the organisation. The extent of damage caused by corrupted data can be greater for systems with high CCP.

Figure 6.2 represents a simple example of the CCP metric. There are five systems, and the blue arrows depict the data flows between these systems. System B sends and receives data from system A, which also sends data to system C. System B also receives data from system D but does not send data to it. The CCP of system B is therefore 3. In the scenario that data in system B was compromised, this would mean that this data could be sent to system A which could then send it to system C. Hence, the corrupted data in system B could be propagated to two other systems and so, three systems would be affected by this data. In the example, systems D and E have the biggest CCP which means that the organisation should closely monitor these two systems, as a data manipulation event in one of these systems would affect all the systems, shutting down the business completely.

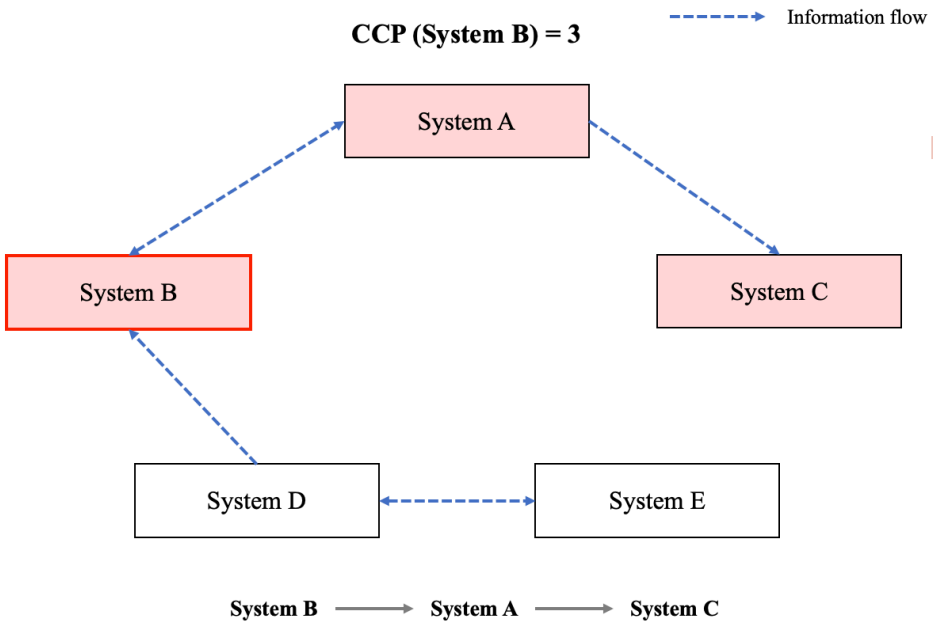

Figure 6.2 – Example of the CCP metric at the system level

This metric can also be used at a lower level, that is at the class level. The CCP metric can also help identify the classes within the systems that will impact others the most in order to better protect and monitor them. In this case, the CCP metric measures the number of classes that could be affected by corrupted data in a specific class. Measuring the CCP of classes can be used as a complement to the systems' CCP. For example, in Figure 6.2, systems A and B both have a CCP of 3, adding the CCP of their classes would help focus the organisation's efforts on specific classes that are more vulnerable to propagating altered data.

Figure 6.3 represents an example of the CCP metric at the class level. The CCP of class A5 within system A is 7. Indeed, class A5 shares data with class C1, A4, and A3 which shares data with class A2, which exchanges data with class A1 which sends data to class B1. In the worst-case scenario, corrupted data in class A5 could affect 7 classes within three different systems.

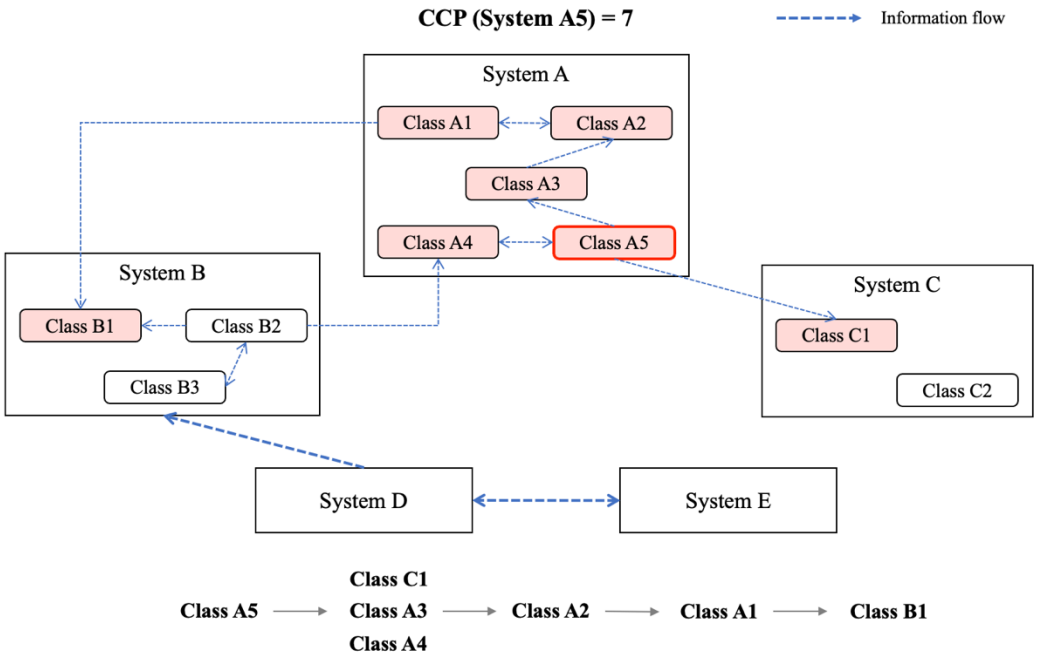

Figure 6.3 – Example of the CCP metric at the class level

Critical Element Ratio (CER)

There is a security risk if some critical elements are changed that can damage the whole system. This metric aims to identify a correlation between the sensitivity of a class and the risk of attacks: the more critical elements in a class, the higher the sensibility of the class, and the higher the risk. This metric represents the number of critical elements in a class divided by the number of elements in the class. It measures how many elements of a class can propagate corrupted data to other classes. The higher the CER of a class is, the more the class is prone to spread compromised data. The CER metric helps increase the protection and monitoring of vulnerable classes as a complement to the CCP metric. For example, two classes can have the same CCP, and measuring their CER will help determine which one contains the most elements likely to propagate corrupted information.

Figure 6.4 depicts a basic example of the CER metric. Class A5 is composed of four elements, three of them are considered critical as they share data with other classes, which means that the CER of class A5 is 0.75. In comparison, the CER of class A4 is 0.66.

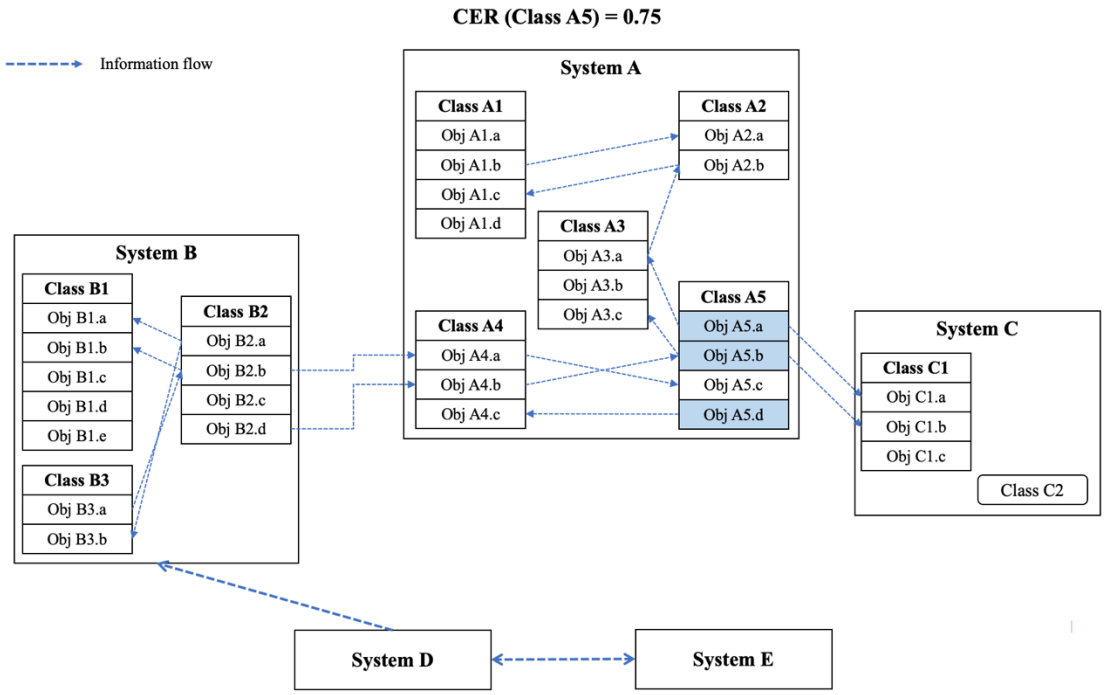

Figure 6.4 – Example of the CER metric

Data validation

Data validation is critical as it aims to validate the accuracy and details of data in order to create the best outcomes possible. If the data is not accurate, then every result or decision based on this data will be inaccurate as well, which could lead to serious consequences. It is necessary to verify and validate data before using it in order to avoid intentional or unintentional errors. Data validation means verifying that the data is valid (i.e., the input is in the expected format and is within an acceptable range) and secure before processing it. To estimate data manipulation (both intentional and unintentional), a security metric could be to measure the percentage of data that have values within the domain of acceptable values (i.e., defined format and range). This metric helps to protect the accuracy of the data by determining if there is a need to improve data verification training and data validation techniques.

6.3.3.2 Access management

Access control

Access control is a security technique that regulates who or what can access or use organisational resources such as data or devices. It helps improve the overall security of the organisation and minimise risks of unauthorised access to an organisation's assets. Access control is the process of managing who is authorised to access data and resources within an organisation by identifying users through various

credentials. Access control metrics help ensure that employees have only access to data and/or systems that are necessary to their work. There are a lot of different metrics but we focus primarily on 1) the number of people who have administrative access, and 2) the distribution of users of a system according to their type of system privileges. Higher is the first metric, and higher is the risk of data manipulation and other cyber threats. The IT team should verify if all the users with administrative privileges still need them in order to adjust their system privileges according to their needs and thus limit the number of vulnerabilities. The second metric helps determine the number of users that can perform specific actions within a system, which can be useful in case of data manipulation attacks.

Account deactivation

Accounts associated with former employees as well as temporary access accounts created for contractors and other third parties that are no longer actively used are a potential way to gain unauthorised access to organisation resources, applications, or systems. Organisations that do not take the necessary efforts to close these entry points leave the door open for malicious actors. Deactivating users' accounts should be a standard procedure, yet it is not uncommon that those accounts are not necessarily deactivated right away. The longer those accounts stay active, the higher the risk that attackers use them to access the organisation's network and sensitive information. Identifying and closing those accounts as soon as possible helps significantly minimise the risks. The number of days it takes to deactivate a former employee, vendor, or contractor user account is a good security measure to determine the average time during which these accounts represent vulnerabilities for the organisation, and it is necessary to keep track of any suspicious activity on these accounts. Furthermore, keeping track of accounts that have been inactive for a certain period of time can help identify accounts that have been forgotten to be deactivated, and thus reduce risks.

6.3.3.3 Monitoring and Response

Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR)

It is critical for organisations to detect and respond to cyber threats as quickly as possible in order to protect themselves from data breaches and cyber-attacks. The rapidity with which an organisation detects and responds to a data breach can have an impact on the damage caused. MTTD and MTTR indicators provide organisations with significant information to identify areas for improvements in their security defence (policies, processes, training). Organisations should quantify and track such indicators to evaluate their performances in terms of security incidents detection and resolution. The mean time to detect is the average time it takes to detect a security incident or system failure while the mean time to respond refers to the average time it takes to recover from a security incident or system failure. The MTTD and MTTR track how effective organisations are at identifying problems and starting the resolution process, and neutralising

problems respectively, as the longer problems stay undetected and unresolved, the more harmful they become. These two metrics help organisations determine areas of improvement in IT training and in the quality of the incident response plan. The MTTD and MTTR can be customised to assess the detection and resolution performance by teams, systems, or types of incidents for example. Besides, other metrics can be defined to complement these two such as mean time to recovery, mean time to repair, or mean time to failure.

Systems with known vulnerabilities

A key cybersecurity metric for measuring the risk that an organisation is facing is the number of vulnerable systems within its environment. Managing updates and patches is a time-consuming process, but it is necessary to avoid security weaknesses in the organisation that could be exploited by malicious actors. Knowing the number of systems with known vulnerabilities as well as what are the vulnerabilities for each system enables an organisation to have a good understanding of its environment, determine the risk exposure of each system, and to better manage risks. This metric also helps organisations adapt their security policy in order to improve the security of these systems, by limiting the number of people accessing them for example. Besides, this metric helps support the development of a vulnerability management plan for these systems as well as test the effectiveness of this plan.

Cybersecurity attacks attempts

All organisations experience security incidents. A good metric is the number of incidents reported by employees within the organisation as it helps determine if employees have assimilated security training and are able to recognise issues. This metric helps track the effectiveness of the organisation's security training. Besides, this metric can be changed to determine the number of incidents reported by monitoring tools over a certain period of time. In this case, the metric helps organisations understand where resources need to be focused in the cybersecurity program.

Moreover, another interesting metric is the number of attack attempts, such as intrusion or data corruption, both successful and unsuccessful, identified with the organisation. This metric aims to evaluate the attempts to attack (intrusion or corruption) the organisation's assets by sorting them by systems or type of attacks for example. This helps improve the security of the system(s) that present the most attempts as well as improve the security training or expand the IT security teams if necessary. Those two metrics aim to assess the efficiency of the security training for identifying suspicious events through the first metric and adopting good practices to reduce security incidents via the second metric.

6.3.3.4 Network metrics

Unidentified devices

Internet of Things devices, visitor devices, or personal devices are regularly added to organisation networks, which increases security risks. As these devices do not necessarily have strong antivirus systems or up-to-date patches, they can introduce accidentally malware and other cyber risks to the network, and ultimately, they pose a serious threat to the organisation. The devices on the network have to be identified in order to understand the threat and the vulnerabilities exposure of the network, which is why network intrusion detection systems are such a crucial aspect of an organisation's security policy implementation. Determining the percentage of unidentified devices on the organisation's network is a metric that helps reduce the risk and security vulnerabilities of the network. Organisations should be able to identify every device on their network, and defining a method to keep track of unidentified devices will help keep this metric as low as possible as well as to better secure the network.

Level of preparedness

All organisations should assess their readiness for any potential cybersecurity threat or attack. Vulnerability scans and vulnerability management tools are essential for reducing vulnerability risk in the network. The level of preparedness represents the number of devices on the organisation network that are fully patched and are up to date with the latest security controls and updates as well as the latest software and antivirus updates. This metric can be determined by executing a vulnerability assessment of the system and regular vulnerability scans. If outdated or not fully-patched devices are detected, these devices would need to be updated and patched. Having devices not fully patched and updated makes them more vulnerable and thus makes the organisation more vulnerable to cyber attacks. Regular vulnerability scans would help keep this metric as low as possible.

Cyber threat hunting

Monitoring network traffic for unusual activity helps the timely detection and containment of cybersecurity threats. Monitoring is used to identify unusual or suspicious activities or behaviours in organisations' systems and networks. Network traffic analysis is a process of constantly monitoring and analysing network traffic in order to define normal traffic patterns as references and trigger incident response if unusual patterns are detected. Network traffic analysis tools can detect unusual activities, which can then be investigated for possible security vulnerabilities. In order to identify suspicious activities on the organisation's network, activities that are potentially designed to compromise the systems, a good security metric is to keep track of the volume of traffic on the whole network. A significant variation of this metric, both lower or higher volume, helps detect and respond to attacks as quickly as possible to limit the damage.

Moreover, keeping track of the average number of CRUD (create, read, update, remove) requests on the database per day can also help detect unusual activities and attempt to compromise, steal, or delete data.

6.3.3.5 Security assessment

Training and awareness

Organisations must understand the human factor in their cybersecurity posture. Cyber attackers are becoming better at exploiting vulnerabilities based on human behaviours to gain access to organisations' networks. End-users are the weakest link in the cybersecurity chain, as they commit simple mistakes like downloading malware or clicking links in phishing emails that often happen to be the major cause of many severe data breaches. Training employees on how to identify suspicious activities or cyber threats is essential for organisations, as it makes them more prepared to prevent data breaches. Many metrics can be used to measure the effectiveness of an organisation's security awareness training. First, it is important to measure the percentage of employees who have completed the training because all employees should be up to date with their training to help understand what the risks are and how they can help protect the organisation’s assets. Second, training often comes with quizzes, and keeping track of the training results is a good practice to ensure that employees understand the training material and to check if there is a need for additional support. These two metrics help evaluate the readiness of organisations’ employees in terms of cybersecurity.

Furthermore, in complement to security awareness training, Security Operations Center (SOC) can organise various campaigns such as phishing campaigns and detection campaigns to test the reaction of employees. Phishing campaigns include simulated phishing attacks, while detection campaigns can involve the injection of false data into the system or databases. Measuring the number of employees that fell for the false phishing attack or the number of false data reported by system or databases’ users as well as cybersecurity tools helps determine the effectiveness of cybersecurity awareness training. These metrics aim to assess the impact of the training on the employees' behaviour and improve the training if necessary.

6.3.4 Information Flow Awareness Training

Cyber-attacks have become more frequent, more sophisticated, and more extensive. Organisations are responding to these threats by investing in security programs and tools. Yet it is not enough as attackers have learned to bypass security controls and gain access to organisations' networks by targeting underprepared employees. A 2014 IBM report stated that human error is a contributing factor in 95 percent of all security incidents investigated [264]. Most commonly human errors are clicking on an infected attachment or unsafe URL, use of default user names and passwords or easy-to-guess passwords, system

misconfiguration, or disclosure of regulated information via the use of an incorrect email address [264]. By exploiting human behaviours, attackers can easily install malware and gain access to sensitive data. As many security risks come from human behaviour, implementing awareness training for employees can be a valuable strategy for preventing, detecting, and reporting security breaches. Cybersecurity awareness training is critical for all employees, regardless of their position in the organisation, as all of them can be targeted by cyber attackers. Security awareness training helps reduce risk, preventing the loss of personal or sensitive information, and money, as well as damage to the organisation's reputation. Effective awareness training programs address the mistakes (e.g., clicking on unsafe URL, improper document disposal, poor patch management) that employees may make that lead to security incidents.

As a result, in addition to the user stories and security metrics, we suggest recommendations for the development of an information flow awareness training that will complement classic security training. Most security awareness trainings focus on phishing attacks, passwords and authentication, physical security (e.g., unattended computers or devices), internet and email use, social engineering, cloud security, and remote work. The objective of the proposed information flow awareness training is to raise awareness about data manipulation attacks as well as to help people identify integrity threats and understand the network vulnerabilities. This information flow awareness training focuses on the representation of information flow within the organisation and between the different business units, and the presentation of the systems' risks. It is designed for employees who are involved with IT systems and have access to the data on a daily basis. It aims to help employees understand the information flow within the organisation, as well as the source and destination of the data that are passing in their business units. Thus, employees will get a better understanding of the damage data manipulation attacks can have on their organisation by demonstrating how far tampered data can spread between the business units and how fast processes can be corrupted. The information flow awareness training we propose is composed of two main parts: 1) raising awareness on information flows and data integrity threats and demonstrating the propagation of corrupted data and its consequences using examples based on the information flows within the organisation; 2) validating the knowledge employees learn in the first part through exercises. This training is customisable to adapt the information flows used as examples to the different business units, and to adjust to the roles and responsibilities of the employees as well as the sensibility of the data they consume. The higher the sensibility of the data, the more advanced the training is.

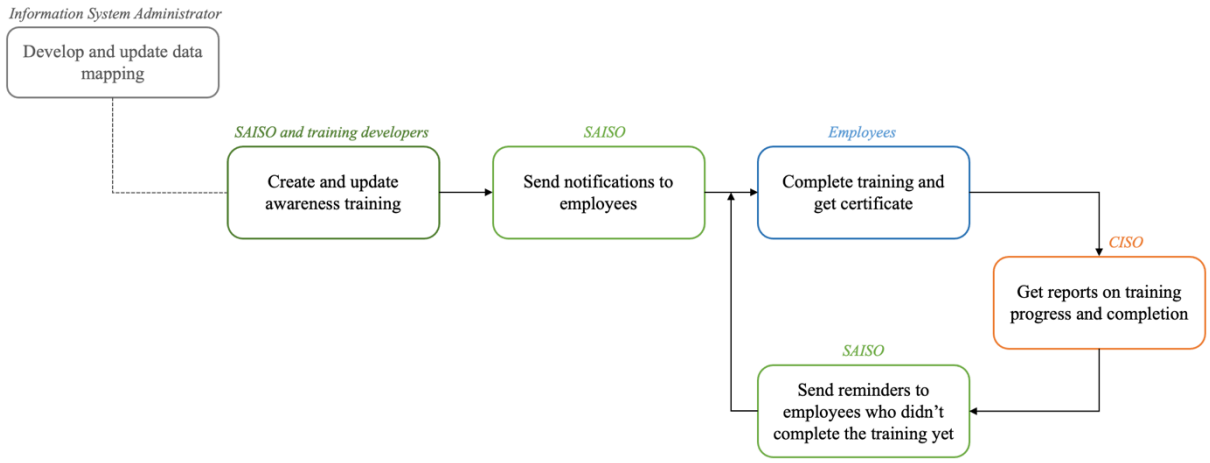

Figure 6.5 – Information Flow Awareness Training

Figure 6.5 depicts recommendations on how this awareness training should be developed. The Senior Information Security Officer works with the training developers to create a comprehensive training using the data mapping developed by the information system administrator, and update it when necessary. Then, the Senior Information Security Officer notifies employees about the training, gives them additional information, and sends regular reminders to employees who have not yet completed the training. Employees take the training and get a certificate (with unique ID and date) once the training is complete. Finally, the Chief information security officer gets reports on employees' training progress and completion. These reports represent for example the distribution of employees according to the status (i.e., to do, in progress, complete) of their training, the distribution of employees by business units according to the status of their training, and the average time taken to complete the training.

6.4 Conclusion

Data manipulation is a cyber-attack aiming to compromise the integrity of data. These attacks can have significant consequences for organisations. Organisations can implement measures to prevent and mitigate data manipulation attacks. These measures help i) identify and protect critical systems and data; ii) monitor networks and detect suspicious activities, and iii) develop responses for mitigating data manipulation events. In this chapter, we propose a Cybersecurity Framework Profile for data manipulation risk management that aligns with the NIST CSF's security objectives. This Profile aims to help organisations identify and prioritise opportunities to improve their data integrity strategy. In addition to this data manipulation Profile, we propose implementation recommendations to help organisations implement a comprehensive solution for data traceability. These recommendations consist of a collection of user stories for the Subcategories of the Profile, a set of security metrics, and guidelines for information flow awareness training.

Industry 4.0 technologies enables organizations to be more connected, which bring new cybersecurity concerns as it creates more entry points that attackers can find and exploit. Organisations are exposed to a growing number of cyber challenges, and as a result, data are also vulnerable to numerous cyber threats, including data manipulation. The growing complexity of the cyber threats led to an increased speed at which tampered data can propagate through organisations and corrupt data exchanges and data-driven decisions. A formal data traceability solution is needed to track data flows across large and complex networks of organisations, and identify and contain the propagation of corrupted data. To address the data traceability challenge, we 1) reviewed existing cybersecurity frameworks to find the framework that would offer a solid and customisable foundation to support our needs for addressing data manipulation risk; 2) proposed a cybersecurity framework for data manipulation risk management, based on the NIST Cybersecurity Framework, to help organisations prevent and mitigate data manipulation events, and finally 3) proposing implementation guidelines (user stories, security metrics, information flow awareness training) to complement the Profile and allowing organisations to develop and implement a comprehensive data traceability solution for data integrity.

CONCLUSION AND FUTURE WORK

The digital transformation of manufacturing offered by Industry 4.0 has resulted in the generation of large volumes of data. Organisations consume and exchange this data while preserving their integrity and quality. Organisations' communication and collaboration rely on fast and reliable data exchanges, but these data exchanges come with data interoperability and data traceability challenges. Using information standards to address these challenges, this thesis proposes a new standard-based framework that 1) presents the transition to adaptive project management and a better requirement elicitation model for the standard development process, and 2) provides a formal data traceability solution through a data manipulation Profile to tackle the cybersecurity issues

Summary

The digital transformation of manufacturing has led to more connected, automated, and data-driven environments and processes. Data has become a key enabler to processes, exchanges, and decision-making. Manufacturing relies heavily on data and the exchange of this data between the different stakeholders, machines, and systems. Data exchange, by definition, is the process of sending and receiving data in a way in which the data content or meaning has not been tampered with during transmission, in other words that the data received is an accurate representation of the data sent.

The digitalisation of manufacturing has emphasised the importance of information management, data exchange, and the interoperability of the different actors in the manufacturing processes. The emergence of new technologies and networked data sources support new opportunities for organisational collaboration through high-speed and high-volume data exchange. In other words, this digital era helped improve the speed, volume, accuracy, and consistency of data exchange and innovations across and within organisations. Data is now available and shareable at a quicker rate than organisations can consume and track it. This speed comes with key challenges. In this thesis, we focused on two of these challenges, data interoperability and data traceability, two interdependent challenges that manufacturers face and must understand to address them.

On one hand, faster innovation and collaboration are being hindered by data interoperability challenges. Increased collaboration is associated with an increased number of heterogeneous systems that need to communicate with each other. Information standards are a proven solution, as they provide a common language to the information systems that support business and operations in organisations. However, their development process is often long and complex because it requires numerous, international, and heterogeneous stakeholders (e.g., users, technical experts, implementers) to collaborate and reach

agreements, oftentimes on a volunteer basis. The standards development process faces inefficiencies that make it incompatible with the fast-paced environment they need to support and provide interoperability for, as well as the ever-evolving nature of the cyber threats that organisations are facing. These inefficiencies therefore slow down standards' publication and adoption. In this thesis, we propose a transition from predictive to adaptive project management with the use of Agile methods. We present some Agile practices, through a test case using STEP AP242, that can help standard development teams overcome some of these inefficiencies. Agile provides a way (through processes and tools) to 1) improve the management of the available resources (i.e., volunteer staff) by reducing the need for long-term planning and commitment, and 2) improve collaboration between these communities by supporting a proper and faster communication channel. Despite the benefits Agile brings, it also comes with challenges. We identified a clear need for better traceability and visibility of requirements, their elicitation, their management, and their implementation. To address this need, we developed and proposed a new requirement elicitation model that serves as a foundation for providing requirements and decision traceability and visibility, by recording and leveraging project meetings in a formal way. We also proposed an implementation of this model through a customisation of Jira, a tool used by the STEP development teams for bug and issue tracking.

On the other hand, Industry 4.0 has enabled manufacturing to become more digitalised, connected, smarter, and more autonomous, leading to a significant increase in the number and complexity of cybersecurity challenges. The digitalisation of Industry 4.0 supports high-volume and high-speed data exchanges across supply chains, enabling a new data-driven approach to business decisions, but also creating a greater attacks surface. These decisions are exposed to the speed at which tampered data can propagate through organisations and corrupt these decisions. Data manipulation can lead to corrupted decisions and processes, as well as compromised entire supply chains, which can have serious consequences for organisations, such as economical and reputational damages. The mean time to identify (MTTI) such a threat is already close to 215 days [17], and the constant growth of data produced and exchanged is likely to push the MTTI upwards. Organisations generally take longer to recover from data manipulation events because 1) unlike ransomware, they take time to be identified, 2) they use data exchange, a key element to industry 4.0, as a propagation vector, and 3) propagation makes it hard to determine the extent and impact of these attacks. While digital signatures have demonstrated their use in identifying such corruption, there is still a need for a formal data traceability framework to track data exchange across large and complex networks of organisations in order to identify and contain the propagation of tampered data. This thesis analysed existing cybersecurity frameworks and their limitations, and proposed a new standard-based framework, in the form of an extended NIST Cybersecurity Framework Profile, to mitigate, manage and track data manipulation attacks. This Profile aims to help organisations align their data integrity prevention and mitigation requirements and resources with the elements of the NIST Cybersecurity Framework, as

well as assess the state of their readiness against such attacks. This Profile comes with implementation recommendations to facilitate its adoption and implementation. These implementation recommendations complement the Profile for developing a formal data traceability solution to detect and mitigate data manipulation events. These recommendations consist of user stories, security metrics and guidelines for information flow awareness training.

Future work

As presented in the previous section, we have focused as a priority on some challenges related to data interoperability and data traceability, but there are still challenges to be addressed in order to have the most complete solution possible to ensure fast and reliable data exchange. The framework we propose in this thesis has some limitations. In the next section, we will discuss these limitations and present solutions on how to address them in the future.

Data interoperability

The Requirement elicitation model we propose in Chapter 4 supports better traceability and visibility of the requirements and decisions. This model is originally designed for the development of information standards, but it can be used by any project with distributed stakeholders. The model presented in this thesis can be extended, that is new attributes (e.g., user story, use case…) can be added to the different concepts, or values of the enumeration lists can be changed, to meet the specific needs of projects. Moreover, this model currently focuses on requirement and decisions gathering and traceability, but it could be extended to support requirement validation and requirement prioritization. For example, a possible enhancement could be to add an additional concept, “Test”, regarding the validation of a requirement or work item. This concept would be intermediate between the Decision and Requirement concepts, and between the Decision and Work Item concepts. This Test concept would be composed of attributes such as acceptance criteria, unit test, integration test, acceptance test or user test for example. This concept would represent the results of the different testing that support the validation of the requirements and/or work items.

The next step for this model is to display the information captured in a meaningful way to the various project stakeholders. Visibility is key to Agile management and project information is often shared graphically in an information radiator that is easily accessible and understandable by all stakeholders. Visual representations of requirements and associated work items help communicate, clarify and validate them, by minimising complexity and misconceptions, as well as identifying gaps in requirements. Identifying the most appropriate visualisation techniques to display the information collected through the model would be the first task in this next step. In addition to these visual requirement models, the model

should also support report generation. Like the visual representation of the requirements, reports provide visibility. Reports help communicate how effectively the project is progressing as well as highlight the challenges and risks the project is facing. These two functionalities would bring significant benefits to project teams, ensuring that the requirements align with the business needs, that all the stakeholders understand the requirements and their context, and that the project is progressing well.

In this thesis, we focus on improving the standards development process to facilitate and promote their adoption and thus address the data interoperability issues. As mentioned, the standards development process is long, and this thesis proposes a solution to shorten the development iterations and increase the delivery velocity of standards development teams, by transitioning to adaptive project management, leveraging Agile methods, and integrating the Requirement elicitation model. The creation and development of a Requirements acceptance framework can also help understand what slows down the standards development process. This framework will use some of the information captured by the Requirement elicitation model. The objective of the Requirement acceptance framework is the analysis of the decision taken during the meetings. This framework will measure and analyse the process of capturing and accepting the requirements by the participants. It will also help project teams to get an overall idea of how the development process is going. Indeed, this framework will evaluate the decisions taken to understand what may cause some disruptions to the project.

Finally, a future step would be to develop long-term storage of standard requirements. The standard development iterations currently take between 18 and 48 months, and standards may be developed through multiple iterations. As a result, it is crucial to ensure that any stakeholders of the standard development can access requirements data during the entire standard development process. An ideal solution would be to store both requirements that have been implemented and requirements that have been rejected in order to maintain traceability of the progress of the industry needs. It would also be beneficial to save information about the context of the requirements and the discussions that have happened about the requirements. In addition, this long-term storage solution will need to ensure the integrity of the requirements data over time.

Data traceability

In this thesis, we focused on data manipulation, which has become a growing threat in cybersecurity. Data manipulation, the concept of subtly modifying data in order to corrupt processes, decisions, or products, can take different forms. As a result, our future work will focus on defining a cybersecurity taxonomy of data manipulation threats. This taxonomy will characterise and classify these threats to allow a precise definition of vulnerabilities that they exploit. Besides, while tools continue to be refined to monitor

and detect data manipulation attacks, educating and training employees about these attacks is also important. Adding this cybersecurity taxonomy to the information flow awareness training we defined in a previous section, will help employees anticipate, recognise, and report suspicious activities and thus help organisations avoid or limit the damage of this kind of attack.

The data manipulation Profile we propose provides organisations with a voluntary, risk-based approach for improving their cybersecurity posture and reducing risks regarding data manipulation. It could be interesting to complete this Profile with standardised vulnerability and asset management tools, which would allow organisations to have better visibility into the most vulnerable parts of their system. Some components of the SCAP framework we present in Section 5.1.2.4, could be used for this purpose. First, for asset management, one could use: Asset Identification (AID) and Asset Reporting Format (ARF) to uniquely identify assets and thus facilitate reporting and correlation of asset information throughout and between organisations; the Common Platform Enumeration (CPE) entity to automate software inventories and enable security compliance checking; finally, the identification schema Software Identification Tagging (SWID) to track the software installed on the devices belonging to the organisation. Second, regarding vulnerability management, it might be useful to add to our solution: Common Vulnerabilities and Exposures (CVE) to identify security vulnerabilities and correlate vulnerability information about software; the Common Vulnerability Scoring System (CVSS) to estimate the severity of a vulnerability and thus prioritise the organisation's responses to known vulnerabilities. All these are only examples of what could be done in the event that we complete our Profile with SCAP.

Finally, the next step will be to develop a tool to implement the features we discussed in Section 6.3.1. This tool will enable organisations to map and monitor the data flows within their supply chains in order to identify risks and vulnerabilities, and thus better protect them. This tool will also enable organisations to catalog their assets (i.e., physical devices, systems, external systems, software, and applications) and inventory the systems of their employees. This tool will also include features to measure and represent the security metrics we defined in Section 6.3.2, as well as features to help create a security training and generate reports based on the recommendations, we made in Section 6.3.3. In addition, this tool should also include the two points discussed previously, that is allowing employees to access the taxonomy on data manipulation threats and incorporating some of SCAP components to help IT teams with vulnerability and asset management. Moreover, to monitor the activities traffic, it would be helpful to define a framework to monitor the flows of HTTP requests, GET and POST, from the different systems as well as the flows of operations (i.e., create, read, update and delete) on the different databases, to determine who is consuming or modifying data. This framework should be integrated into the tool to help with the monitoring of the data flows.

Data is a key asset for organisations as it plays a crucial role in organisational decision-making and strategy. Data exchanges need to be accurate and understandable by all organisational systems and stakeholders for decisions and processes based on these exchanges to be reliable and trustworthy. Therefore, in this thesis, we addressed two main and interdependent challenges related to data exchange: data interoperability and data traceability. This thesis focused on facilitating data exchange while preventing data manipulation through the use of information standards.

To conclude, this thesis aimed to reduce the time of 1) development and implementation of data interoperability solutions and 2) data traceability operations in response to cyber-attacks that manufacturers are victims of, in order to meet the speed at which data exchanges are now set up and performed.

REFERENCES

[1] Reinsel, D., Gantz, J., & Rydning, J. (2018). *The digitization of the world from edge to core*. IDC. https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

[2] Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2016). *Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries*.

[3] Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*, *48*. https://doi.org/https://doi.org/10.1016/j.jmsy.2018.01.006.

[4] Panetto, H., Iung, B., Ivanov, D., Weichhart, G., & Wang, X. (2019). Challenges for the cyber-physical manufacturing enterprises of the future. Annual Reviews in Control, 47, 200–213. https://doi.org/10.1016/j.arcontrol.2019.02.002

[5] Panetto, H. (2007). Towards a Classification Framework for Interoperability of Enterprise Applications. *International Journal of Computer Integrated Manufacturing*, Taylor & Francis, 2007, 20 (8), pp.727-740. https://doi.org/10.1080/09511920600996419

[6] Sapp, B., Harvey, M., Toussaint, M., Krima, S., Barnard Feeney, A., & Panetto, H. (2021). *Agile for model-based-standards development*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.ams.100-40

[7] Agostinho, C., Ducq, Y., Zacharewicz, G., Sarraipa, J., Lampathaki, F., Poler, R., & Jardim-Goncalves, R. (2016). Towards a sustainable interoperability in networked enterprise information systems: Trends of knowledge and model-driven technology. *Computers in Industry*, *79*, 64–76. https://doi.org/10.1016/j.compind.2015.07.001

[8] ISO. (2020). *ISO 10303-242:2020*. ISO. https://www.iso.org/standard/66654.html

[9] ISO. (2009). *ISO 6983-1:2009*. ISO. https://www.iso.org/standard/34608.html

[10] ISO. (2007b). *ISO 10303-238:2007*. ISO. https://www.iso.org/standard/38036.html

[11] MTConnect Institute. (n.d.). *MTConnect*. MTConnect. Retrieved April 25, 2022, from https://www.mtconnect.org/

[12] DMSC. (2016). *Digital metrology standards consortium QIF & DMIS standards*. DMSC. https://qifstandards.org/

[13] ISO. (n.d.). *Stages and resources for standards development*. ISO. https://www.iso.org/stages-and-resources-for-standards-development.html

[14] inray Industriesoftware GmbH. (2018). Industry 4.0 depends on data transfer. Timely and target-oriented. *Medium*. https://medium.com/@inray.industriesoftware/industry-4-0-depends-on-data-transfer-timely-and-target-oriented-e725eeffa3e3

[15] IBM. (2022). IBM: X-Force threat intelligence index. *Computer Fraud & Security*, *2022*(3). https://doi.org/10.12968/s1361-3723(22)70561-1

[16] Ham, J. V. D. (2021). Toward a better understanding of "cybersecurity." *Digital Threats: Research and Practice*, *2*(3), 1–3. https://doi.org/10.1145/3442445

[17] Nweke, L. (2017). Using the CIA and AAA Models to explain Cybersecurity Activities. *PM World Journal*, *VI*(XII).

[18] IBM. (2021). IBM: Cost of a data breach report. *Computer Fraud & Security*, *2021*(8), 4–4. https://doi.org/10.1016/s1361-3723(21)00082-8

[19] Wu, D., Ren, A., Zhang, W., Fan, F., Liu, P., Fu, X., & Terpenny, J. (2018). Cybersecurity for digital manufacturing. *Journal of Manufacturing Systems*, *48*, 3–12. https://doi.org/10.1016/j.jmsy.2018.03.006

[20] IBM. (2021). IBM: 2021 X-force threat intelligence index. *Network Security*, *2021*(3), 4–4. https://doi.org/10.1016/s1353-4858(21)00026-x

[21] Agarwal, A., & Agarwal, A. (2011). The Security Risks Associated with Cloud Computing. *International Journal of Computer Applications in Engineering Sciences*, *1*.

[22] Hedberg, T. D., Jr., Krima, S., & Camelio, J. A. (2016). Embedding X.509 Digital certificates in three-dimensional models for authentication, authorization, and traceability of product data. *Journal of Computing and Information Science in Engineering*, *17*(1). https://doi.org/10.1115/1.4034131

[23] ISO. (n.d.). *Stages and resources for standards development*. ISO. Retrieved February 2, 2022, from https://www.iso.org/stages-and-resources-for-standards-development.html

[24] Pratt, M. J. (2005). ISO 10303, the STEP standard for product data exchange, and its PLM capabilities. *International Journal of Product Lifecycle Management*, *1*(1), 86. https://doi.org/10.1504/ijplm.2005.007347

[25] ISO. (2004). *Industrial automation systems and integration — Product data representation and exchange — Part 11: Description methods: The EXPRESS language reference manual (ISO 10303-11:2004)*. ISO. https://www.iso.org/standard/38047.html

[26] ISO. (2002). *Industrial automation systems and integration — Product data representation and exchange — Part 21: Implementation methods: Clear text encoding of the exchange structure (ISO 10303-21:2002)*. ISO. https://www.iso.org/standard/33713.html

[27] ISO. (2007). *Industrial automation systems and integration — Product data representation and exchange — Part 28: Implementation methods: XML representations of EXPRESS schemas and data, using XML schemas (ISO 10303-28:2007)*. ISO. https://www.iso.org/standard/40646.html

[28] Barnard Feeney , A. (2002). The STEP modular architecture. *Journal of Computing and Information Science in Engineering*, *2*(2), 132–135. https://doi.org/10.1115/1.1511520

[29] Jardim-Gocalves, R., Olavo, R., & Steiger-Garcao, A. (2005). The emerging ISO10303 Modular Architecture. *International Journal of IT Standards and Standardization Research*, *3*(2), 82–95. https://doi.org/10.4018/jitsr.2005070107

[30] Edeki, C. (2015). Agile software development methodology. *European Journal of Mathematics and Computer Science*, *2*(1).

[31] Kumar, G., & Bhatia, P. (2012). Impact of Agile Methodology on Software Development Process. *International Journal of Computer Technology and Electronics Engineering (IJCTEE)*, *2*(4).

[32] Mabrouk, A., Penas, O., Plateaux, R., Barkallah, M., Choley, J.-Y., & Akrout, A. (2018, October). Integration of agility in a MBSE methodology for multidisciplinary systems design. *2018 IEEE International Systems Engineering Symposium (ISSE)*. http://dx.doi.org/10.1109/syseng.2018.8544425

[33] Shameem, M., Chandra, B256., Kumar, C., & Khan, A. (2019). Impact of requirements volatility and flexible management on GSD project success: A study based on the dimensions of requirements volatility. *International Journal of Agile Systems and Management*, *12*(3). https://doi.org/10.1504/ijasm.2019.10018758

[34] Thummadi, B. V., Shiv, O., & Lyytinen, K. (2011). Enacted routines in agile and waterfall processes. *2011 AGILE Conference*. http://dx.doi.org/10.1109/agile.2011.29

[35] Fagarasan, C., Popa, O., Pisla, A., & Cristea, C. (2021). Agile, waterfall and iterative approach in information technology projects. *IOP Conference Series: Materials Science and Engineering*, *1169*(1). https://doi.org/10.1088/1757-899x/1169/1/012025

[36] The Standish Group. (2020). *CHAOS Report 2020*.

[37] Donaldson, N. (2018). *How Agile transparency reduces project risk*. Bigger Impact. https://www.boost.co.nz/blog/2018/11/agile-transparency-reduces-project-risk

[38] Digital.ai. (2021). *15th State of Agile Report*.

[39] Rodríguez, P., Markkula, J., Oivo, M., & Turula, K. (2012). Survey on agile and lean usage in finnish software industry. *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement - ESEM '12*. http://dx.doi.org/10.1145/2372251.2372275

[40] Rico, D. F. (2008). What is the ROI of agile vs. traditional methods? An analysis of XP, TDD, Pair Programming, and Scrum (Using Real Options). *TickIT International*, *10*(4), 9–18.

[41] Hunt, J. (2006). Agile methods and the agile manifesto. In *Agile Software Construction* (pp. 9–30). Springer London. http://dx.doi.org/10.1007/1-84628-262-4_2

[42] Stellman, A., & Greene, J. (2014). *Learning agile: Understanding scrum, XP, lean, and kanban*. O'Reilly Media, Inc.

[43] Project Management Institute. (2017). *Agile Practice Guide*.

[44] Scaled Agile. (2017). *SAFe 5.0 Framework*. Scaled Agile Framework. https://www.scaledagileframework.com

[45] PMI. (n.d.). *Disciplined agile delivery*. Disciplined Agile. Retrieved April 25, 2022, from https://www.pmi.org/disciplined-agile/process/introduction-to-dad

[46] Larman, C., & Vodde, B. (2016). *Large-Scale scrum: More with less*. Addison-Wesley Professional.

[47] Agilest. (2016). *Scrum of scrums*. Guide to Agile Scaling Frameworks - Agilest®. https://www.agilest.org/scaled-agile/scrum-of-scrums/

[48] KnowledgeHut. (2018). *LeSS Vs SAFe: Explore your best fit to lift productivity in your team*. KnowledgeHut. https://www.knowledgehut.com/blog/agile/less-vs-safe-which-certification-should-you-choose-and-why

[49] Piikkila, J. (n.d.). *What is SAFe?* Atlassian. Retrieved April 30, 2022, from https://www.atlassian.com/agile/agile-at-scale/what-is-safe

[50] STEP AP242. (n.d.). *Introduction*. EPLM Interoperability. Retrieved April 25, 2022, from http://www.ap242.org

[51] Mohammed, S. K., Arbo, M. H., & Tingelstad, L. (2021). Leveraging model based definition and STEP AP242 in task specification for robotic assembly. *Procedia CIRP*, *97*, 92–97. https://doi.org/10.1016/j.procir.2020.05.209

[52] Wardhani, R., & Xu, X. (2016). Model-based manufacturing based on STEP AP242. *2016 12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*. http://dx.doi.org/10.1109/mesa.2016.7587187

[53] Datakit. (n.d.). *STEP AP242 managed model based 3D engineering*. Datakit. Retrieved April 25, 2022, from https://www.datakit.com/fr/step_ap242.php

[54] Brangé, J., Delaunay, J.-Y., & Moura, E. (2014). *Whitepaper Development of STEP AP 242 ed2 "Managed Model Based 3D Engineering."*

[55] Brangé, J., & Ungerer, M. (2019). Industry Requirements for ISO 10303 5 Years Roadmap – The Way Towards STEP AP 242 Edition 3. In *Prostep*. https://www.prostep.org/fileadmin/user_upload/19_Industry_requirements_for_AP242_5_years_Roadmap_including_AP24_ed3.pdf

[56] Sedano, T., Ralph, P., & Peraire, C. (2019). The product backlog. *2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE)*. http://dx.doi.org/10.1109/icse.2019.00036

[57] Scaled Agile Framework. (2017). *Agile release train*. Scaled Agile Framework. https://www.scaledagileframework.com/agile-release-train/

[58] Van Amerongen, R. (2016). *Agile best agile estimation techniques – beyond scrum planning poker - AMIS, data driven blog - Oracle & microsoft azure*. AMIS Technology Blog. https://technology.amis.nl/agile/8-agile-estimation-techniques-beyond-planning-poker/

[59] Chen, L. (2017). Continuous Delivery: Overcoming adoption challenges. *Journal of Systems and Software*, *128*, 72–86. https://doi.org/10.1016/j.jss.2017.02.013

[60] Chen, L. (2015). Towards architecting for continuous delivery. *2015 12th Working IEEE/IFIP Conference on Software Architecture*. http://dx.doi.org/10.1109/wicsa.2015.23

[61] Scaled Agile Framework. (2017c). *Continuous exploration*. Scaled Agile Framework. https://www.scaledagileframework.com/continuous-exploration/

[62] Scaled Agile Framework. (2014). *Continuous integration*. Scaled Agile Framework. https://www.scaledagileframework.com/continuous-integration/

[63] Stolberg, S. (2009). Enabling agile testing through continuous integration. *2009 Agile Conference*. http://dx.doi.org/10.1109/agile.2009.16

[64] Scaled Agile Framework. (2017). *Continuous deployment*. Scaled Agile Framework. https://www.scaledagileframework.com/continuous-deployment/

[65] Tassey, G. (2002). *The Economic Impacts of Inadequate Infrastructure for Software Testing Final Report Prepared for*.

[66] Benmoshe, I. (2014). *Whitepaper How to Calculate the ROI of Continuous Delivery*. Zend. http://downloads.zend.com/whitepapers/WP-How-to-Calculate-the-ROI-of-Continuous-Delivery-2014-04-10-EN.pdf

[67] Hilton, M., Tunnell, T., Huang, K., Marinov, D., & Dig, D. (2016). Usage, costs, and benefits of continuous integration in open-source projects. *Proceedings of the 31st IEEE/ACM International Conference on Automated Software Engineering*. http://dx.doi.org/10.1145/2970276.2970358

[68] Schön, E.-M., Thomaschewski, J., & Escalona, M. J. (2017). Agile Requirements Engineering: A systematic literature review. *Computer Standards & Interfaces*, *49*, 79–91. https://doi.org/10.1016/j.csi.2016.08.011

[69] Wiegers, K. E. (2000). Karl Wiegers describes 10 requirements traps to avoid. *Software Testing & Quality Engineering*, *2*(1).

[70] Kumar, V. S. (2006). Effective requirements management. *PMI Global Congress 2006*.

[71] Nguyen, L., & Swatman, P. A. (2003). Managing the requirements engineering process. *Requirements Engineering*, *8*(1), 55–68. https://doi.org/10.1007/s00766-002-0136-y

[72] Rehman, T., Khan, M., & Riaz, N. (2013). Analysis of Requirement Engineering Processes, Tools/Techniques and Methodologies. *International Journal of Information Technology and Computer Science*, *5*, 40–48. https://doi.org/10.5815/ijitcs.2013.03.05

[73] The Standish Group. (2014). *CHAOS Report 2014*.

[74] Paetsch, F., Eberlein, A., & Maurer, F. (2003). Requirements engineering and agile software development. *Proceedings of the IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*. https://doi.org/10.1109/ENABL.2003.1231428

[75] Pandey, D., & Pandey, V. (2012). Importance of Requirement Management : A Requirement Engineering Concern. *International Journal of Research and Development - A Management Review (IJRDMR)*, *1*.

[76] Sharma, S., & Pandey, S. K. (2013). Revisiting requirements elicitation techniques. *International Journal of Computer Applications*, *75*(12). https://doi.org/10.5120/13166-0889

[77] Medeiros, J., Alves, D. C. P., Vasconcelos, A., Silva, C., & Garcia Wanderley, E. (2015). Requirements engineering in agile projects: A systematic mapping based in evidences of industry. *XVIII Ibero-American Conference on Software Engineering (CIBSE)/ESELAW*.

[78] Kauppinen, M., Vartiainen, M., Kontio, J., Kujala, S., & Sulonen, R. (2004). Implementing requirements engineering processes throughout organizations: Success factors and challenges. *Information and Software Technology*, *46*(14), 937–953. https://doi.org/10.1016/j.infsof.2004.04.002

[79] Sommerville, I. (2005). Integrated requirements engineering: A tutorial. *IEEE Software*, *22*(1), 16–23. https://doi.org/10.1109/ms.2005.13

[80] Sadiq, M., & Jain, S. (2012). An Insight into Requirements Engineering Processes. *International Conference on Advances in Communication, Network, and Computing*, *108*. https://doi.org/10.1007/978-3-642-35615-5_48

[81] Sommerville, Ian, & Sawyer, P. (1997). Viewpoints: principles, problems and a practical approach to requirements engineering. *Annals of Software Engineering*, *3*, 101–130. https://doi.org/10.1023/A:1018946223345

[82] Attarha, M., & Modiri, N. (2011). Focusing on the importance and the role of requirements engineering. *The 4th International Conference on Interaction Sciences*

[83] Inayat, I., Moraes, L., Daneva, M., & Salim, S. S. (2015). A Reflection on Agile Requirements Engineering: Solutions Brought and Challenges Posed. *Scientific Workshop Proceedings of the XP2015*. http://dx.doi.org/10.1145/2764979.2764985

[84] Pandey, D., Suman, U., & Ramani, A. K. (2010). An effective requirement engineering process model for software development and requirements management. *2010 International Conference on Advances in Recent Technologies in Communication and Computing*. http://dx.doi.org/10.1109/artcom.2010.24

[85] Cheng, B. H. C., & Atlee, J. M. (2007). Research directions in requirements engineering. *Future of Software Engineering (FOSE '07)*. http://dx.doi.org/10.1109/fose.2007.17

[86] Oberg, R., Probasco, L., & Ericsson, M. (2000). Applying requirements management with use cases. *Rational Software White Paper*

[87] Gambo, I. P., Soriyan, A. H., & Ikono, R. N. (2014). A Proposed Process Model for Requirements Engineering using Delphi Techniques for Prioritisation. *International Journal of Information Technology and Computer Science*, *7*(1), 73–80. https://doi.org/10.5815/ijitcs.2015.01.09

[88] Lehtola, L., Kauppinen, M., & Kujala, S. (2004). Requirements prioritization challenges in practice. In *Product Focused Software Process Improvement* (pp. 497–508). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-540-24659-6_36

[89] Hussain, A., & Mkpojiogu, E. O. C. (2016). Requirements: Towards an understanding on why software projects fail. *AIP Conference Proceedings*. http://dx.doi.org/10.1063/1.4960886

[90] Naz, H., & Khokhar, M. N. (2009). Critical Requirements Engineering Issues and their Solution. *2009 International Conference on Computer Modeling and Simulation*. http://dx.doi.org/10.1109/iccms.2009.50

[91] Bjarnason, E., Wnuk, K., & Regnell, B. (2011). A case study on benefits and side-effects of agile practices in large-scale requirements engineering. *Proceedings of the 1st Workshop on Agile Requirements Engineering - AREW '11*. http://dx.doi.org/10.1145/2068783.2068786

[92] Yaseen, M., Baseer, S., & Sherin, S. (2015). Critical challenges for requirement implementation in context of global software development: A systematic literature review. *2015 International Conference on Open Source Systems & Technologies (ICOSST)*. http://dx.doi.org/10.1109/icosst.2015.7396413

[93] Hoffmann, M., Kuhn, N., Bittner, M., & Weber, M. (2004). Requirements for requirements management tools. *Proceedings. 12th IEEE International Requirements Engineering Conference, 2004.* http://dx.doi.org/10.1109/icre.2004.1335687

[94] Gotel, O., & Mader, P. (2009). How to select a requirements management tool: Initial steps. *2009 17th IEEE International Requirements Engineering Conference*. http://dx.doi.org/10.1109/re.2009.49

[95] Sinha, V., Sengupta, B., & Chandra, S. (2006). Enabling collaboration in distributed requirements management. *IEEE Software*, *23*(5), 52–61. https://doi.org/10.1109/ms.2006.123

[96] Smite, D. (2006). Requirements management in distributed projects. *Journal of Universal Knowledge Management*, *1*.

[97] Almefelt, L., Berglund, F., Nilsson, P., & Malmqvist, J. (2006). Requirements management in practice: Findings from an empirical study in the automotive industry. *Research in Engineering Design*, *17*(3), 113–134. https://doi.org/10.1007/s00163-006-0023-5

[98] Kelanti, M., Hyysalo, J., Välimäki, A., Kuvaja, P., & Oivo, M. (2013). A Case Study of Requirements Management: Toward Transparency in Requirements Management Tools. *The Eighth International Conference on Software Engineering Advances (ICSEA 2013)*.

[99] Knauss, E., Yussuf, A., Blincoe, K., Damian, D., & Knauss, A. (2016). Continuous clarification and emergent requirements flows in open-commercial software ecosystems. *Requirements Engineering*, *23*(1), 97–117. https://doi.org/10.1007/s00766-016-0259-1

[100] Damian, D. E., & Zowghi, D. (2003). RE challenges in multi-site software development organisations. *Requirements Engineering*, *8*(3), 149–160. https://doi.org/10.1007/s00766-003-0173-1

[101] Hafeez, M. S., Rasheed, F., & Mr, K. (2017). An improved model for requirement management system. *Journal of Information Technology & Software Engineering*, *07*(01). https://doi.org/10.4172/2165-7866.1000196

[102] Maeder, P., Riebisch, M., & Philippow, I. (2006). Traceability for managing evolutionary change. *Proceedings of the 15th International Conference on Software Engineering and Data Engineering (SEDE)*.

[103] Sankhwar, S., Singh, V., & Pandey, D. (2014). Requirement engineering paradigm. *Global Journal of Multidisciplinary Studies*, *3*(3).

[104] Violante, M. G., & Vezzetti, E. (2014). A methodology for supporting requirement management tools (RMt) design in the PLM scenario: An user-based strategy. *Computers in Industry*, *65*(7), 1065–1075. https://doi.org/10.1016/j.compind.2014.05.001

[105] Ahmed, H., Hussain, A., & Baharom, F. (2016). Current Challenges of Requirement Change Management. *Journal of Telecommunication, Electronic and Computer Engineering*, *8*, 173–176.

[106] Goknil, A., Kurtev, I., van den Berg, K., & Spijkerman, W. (2014). Change impact analysis for requirements: A metamodeling approach. *Information and Software Technology*, *56*(8), 950–972. https://doi.org/10.1016/j.infsof.2014.03.002

[107] Mustafa, N. (2018). *Traceability modeling for the engineering of heterogeneous systems* [Carleton University]. http://dx.doi.org/10.22215/etd/2018-13442

[108] O'Neal, J. S., & Carver, D. L. (2001). Analyzing the impact of changing requirements. *Proceedings IEEE International Conference on Software Maintenance (ICSM 2001)*. http://dx.doi.org/10.1109/icsm.2001.972729

[109] de la Vara, J. L., Borg, M., Wnuk, K., & Moonen, L. (2016). An industrial survey of safety evidence change impact analysis practice. *IEEE Transactions on Software Engineering*, *42*(12), 1095–1117. https://doi.org/10.1109/tse.2016.2553032

[110] Ramesh, B., Powers, T., Stubbs, C., & Edwards, M. (1995). Implementing requirements traceability: A case study. *Proceedings of 1995 IEEE International Symposium on Requirements Engineering (RE'95)*. http://dx.doi.org/10.1109/isre.1995.512549

[111] Jarke, M. (1998). Requirements tracing. *Communications of the ACM*, *41*(12), 32–36. https://doi.org/https://doi.org/10.1145/290133.290145

[112] Gotel, O. C. Z., & Finkelstein, C. W. (1994). An analysis of the requirements traceability problem. *Proceedings of IEEE International Conference on Requirements Engineering*. http://dx.doi.org/10.1109/icre.1994.292398

[113] Letelier, P. (2002). A framework for requirements traceability in UML-based projects. Proceedings of The1st International Workshop on Traceability in Emerging Forms of Software Engineering.

[114] Pohl, K. (1996). PRO-ART: Enabling requirements pre-traceability. *Proceedings of the Second International Conference on Requirements Engineering*. http://dx.doi.org/10.1109/icre.1996.491432

[115] Khursheed, F., & Suaib, M. (2015). A Survey on Importance of Requirement Traceability in Software Engineering. *International Journal of Engineering and Innovative Technology (IJEIT)*, *5*(4).

[116] Imtiaz, S., Ikram, N., & Imtiaz, S. (2008). Impact analysis from multiple perspectives: Evaluation of traceability techniques. *2008 The Third International Conference on Software Engineering Advances*. http://dx.doi.org/10.1109/icsea.2008.50

[117] Garcia, J. E., & Paiva, A. C. R. (2016). A requirements-to-implementation mapping tool for requirements traceability. *Journal of Software*, *11*(2), 193–200. https://doi.org/10.17706/jsw.11.2.193-200

[118] Ramesh, Balasubramaniam. (1998). Factors influencing requirements traceability practice. *Communications of the ACM*, *41*(12), 37–44. https://doi.org/10.1145/290133.290147

[119] Murugappan, S., Prabha, D., & Student, P. G. (2017). Requirement Traceability for Software Development Lifecycle. *International Journal of Scientific & Engineering Research*, *8*(5).

[120] Torkar, R., Gorschek, T., Feldt, R., Svahnberg, M., Raja, U. A., & Kamran, K. (2012). Requirements traceability: A systematic review and industry case study. *International Journal of Software Engineering and Knowledge Engineering*, *22*(03), 385–433. https://doi.org/10.1142/s021819401250009x

[121] Saiedian, H., Kannenberg, A., & Morozov, S. (2011). A streamlined, cost-effective database approach to manage requirements traceability. *Software Quality Journal*, *21*(1), 23–38. https://doi.org/10.1007/s11219-011-9166-3

[122] Saleem, M., & Minhas, N. M. (2018). Information retrieval based requirement traceability recovery approaches - a systematic literature review. *University of Sindh Journal of Information and Communication Technology (USJICT)*, *2*(4), 180–188.

[123] Spanoudakis, G., Zisman, A., Pérez-Miñana, E., & Krause, P. (2004). Rule-based generation of requirements traceability relations. *Journal of Systems and Software*, *72*(2), 105–127. https://doi.org/10.1016/s0164-1212(03)00242-5

[124] Elghariani, K., & Kama, N. (2016). Review on Agile requirements engineering challenges. *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*. http://dx.doi.org/10.1109/iccoins.2016.7783267

[125] Cao, L., & Ramesh, B. (2008). Agile requirements engineering practices: An empirical study. *IEEE Software*, *25*(1), 60–67. https://doi.org/10.1109/ms.2008.1

[126] Inayat, I., Salim, S. S., Marczak, S., Daneva, M., & Shamshirband, S. (2015). A systematic literature review on agile requirements engineering practices and challenges. *Computers in Human Behavior*, *51*, 915–929. https://doi.org/10.1016/j.chb.2014.10.046

[127] Heikkilä, V. T., Paasivaara, M., Lasssenius, C., Damian, D., & Engblom, C. (2017). Managing the requirements flow from strategy to release in large-scale agile development: A case study at Ericsson. *Empirical Software Engineering*, *22*(6), 2892–2936. https://doi.org/10.1007/s10664-016-9491-z

[128] Alhazmi, A., & Huang, S. (2020). Survey on differences of requirements engineering for traditional and agile development processes. *2020 SoutheastCon*. http://dx.doi.org/10.1109/southeastcon44009.2020.9397492

[129] Bjarnason, E., Wnuk, K., & Regnell, B. (2011). A case study on benefits and side-effects of agile practices in large-scale requirements engineering. *Proceedings of the 1st Workshop on Agile Requirements Engineering - AREW '11*. http://dx.doi.org/10.1145/2068783.2068786

[130] Ramesh, Balasubramaniam, Cao, L., & Baskerville, R. (2010). Agile requirements engineering practices and challenges: An empirical study. *Information Systems Journal*, *20*(5), 449–480. https://doi.org/10.1111/j.1365-2575.2007.00259.x

[131] Batool, A., Motla, Y. H., Hamid, B., Asghar, S., Riaz, M., Mukhtar, M., & Ahmed, M. (2013). Comparative study of traditional requirement engineering and Agile requirement engineering. *2013 15th International Conference on Advanced Communications Technology (ICACT)*.

[132] Alam, S., Bhatti, S. N., Shah, S. A. A., & Jadi, Dr. A. M. (2017). Impact and Challenges of Requirement Engineering in Agile Methodologies: A Systematic Review. *International Journal of Advanced Computer Science and Applications*, *8*(4).

[133] Ahmad, G., Soomro, T., & Naqvi, S. M. R. (2016). An overview: Merits of agile project management over traditional project management in software development. *Journal of Information & Communication Technology*, *10*(1), 105–120.

[134] Lucia, A. D., & Qusef, A. (2010). Requirements Engineering in Agile Software Development. *Journal of Emerging Technologies in Web Intelligence*, *2*(3).

[135] Espinoza, A., & Garbajosa, J. (2011). A study to support agile methods more effectively through traceability. *Innovations in Systems and Software Engineering*, *7*(1), 53–69. https://doi.org/10.1007/s11334-011-0144-5

[136] Cleland-Huang, J. (2011). Traceability in agile projects. In *Software and Systems Traceability* (pp. 265–275). Springer London. http://dx.doi.org/10.1007/978-1-4471-2239-5_12

[137] Lee, C., Guadagno, L., & Jia, X. (2003). An Agile Approach to Capturing Requirements and Traceability. *Proceedings of the 2nd International Workshop on Traceability in Emerging Forms of Software Engineering*.

[138] Taromirad, M., & Paige, R. F. (2012). Agile requirements traceability using domain-specific modelling languages. *Proceedings of the 2012 Extreme Modeling Workshop on - XM '12*. http://dx.doi.org/10.1145/2467307.2467316

[139]. Ernst, N. A., Borgida, A., Jureta, I. J., & Mylopoulos, J. (2014). Agile requirements engineering via paraconsistent reasoning. *Information Systems*, *43*(C), 100–116. https://doi.org/10.1016/j.is.2013.05.008

[140] Heikkila, V. T., Damian, D., Lassenius, C., & Paasivaara, M. (2015). A mapping study on requirements engineering in agile software development. *2015 41st Euromicro Conference on Software Engineering and Advanced Applications*. http://dx.doi.org/10.1109/seaa.2015.70

[141] Rasheed, A., Zafar, B., Shehryar, T., Aslam, N. A., Sajid, M., Ali, N., Dar, S. H., & Khalid, S. (2021). Requirement Engineering Challenges in Agile Software Development. *Mathematical Problems in Engineering*. https://doi.org/https://doi.org/10.1155/2021/6696695

[142] Besrour, S., Rahim, L. B. A., & Dominic, P. D. D. (2016). A quantitative study to identify critical requirement engineering challenges in the context of small and medium software enterprise. *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*. http://dx.doi.org/10.1109/iccoins.2016.7783284

[143] Shah, T., & Patel, S. V. (2014). A review of requirement engineering issues and challenges in various software development methods. *International Journal of Computer Applications*, *99*(15), 36–45. https://doi.org/10.5120/17451-8370

[144] Davis, C. J., Fuller, R. M., Tremblay, M. C., & Berndt, D. J. (2006). Communication challenges in requirements elicitation and the use of the repertory grid technique. *Journal of Computer Information Systems*, *46*(5), 78–86. https://doi.org/10.1080/08874417.2006.11645926

[145] In, H., & Roy, S. (2001). Visualization issues for software requirements negotiation. *25th Annual International Computer Software and Applications Conference. COMPSAC 2001*. http://dx.doi.org/10.1109/cmpsac.2001.960592

[146] Decker, B., Ras, E., Rech, J., Jaubert, P., & Rieth, M. (2007). Wiki-Based stakeholder participation in requirements engineering. *IEEE Software*, *24*(2), 28–35. https://doi.org/10.1109/ms.2007.60

[147] PMI. (2017a). *A guide to the project management body of knowledge (PMBOK guide)* (6th ed.). Project Management Institute

[148] Hadar, I., Soffer, P., & Kenzi, K. (2012). The role of domain knowledge in requirements elicitation via interviews: An exploratory study. *Requirements Engineering*, *19*(2), 143–159. https://doi.org/10.1007/s00766-012-0163-2

[149] Konaté, J., Sahraoui, A. E. K., & Kolfschoten, G. L. (2013). Collaborative requirements elicitation: A process-centred approach. *Group Decision and Negotiation*, *23*(4), 847–877. https://doi.org/10.1007/s10726-013-9350-x

[150] Spoletini, P., & Ferrari, A. (2017). Requirements elicitation: A look at the future through the lenses of the past. *2017 IEEE 25th International Requirements Engineering Conference (RE)*. http://dx.doi.org/10.1109/re.2017.35

[151] Fernandes, J., Duarte, D., Ribeiro, C., Farinha, C., Pereira, J. M., & Silva, M. M. da. (2012). iThink: A Game-Based Approach Towards Improving Collaboration and Participation in Requirement Elicitation. *Procedia Computer Science*, *15*, 66–77. https://doi.org/10.1016/j.procs.2012.10.059

[152] Nahar, N. G., Wora, P. K., & Kumaresh, S. (2013). Managing requirement elicitation issues using step-wise refinement model. *International Journal of Advanced Studies in Computers, Science and Engineering*, *2*(5), 27–33.

[153] Bani-Salameh, H., & Al jawabreh, N. (2015). Towards a comprehensive survey of the requirements elicitation process improvements. *Proceedings of the International Conference on Intelligent Information Processing, Security and Advanced Communication*. http://dx.doi.org/10.1145/2816839.2816872

[154] Adedjouma, M., Dubois, H., & Terrier, F. (2011). Requirements exchange: From specification documents to models. *2011 16th IEEE International Conference on Engineering of Complex Computer Systems*. http://dx.doi.org/10.1109/iceccs.2011.42

[155] Graf, A., & Jastram, M. (2011). *Requirements, Traceability and DSLs in Eclipse with the Requirements Interchange Format (RIF/ReqIF)*.

[156] Ebert, C., & Jastram, M. (2012). ReqIF: Seamless requirements interchange format between business partners. *IEEE Software*, *29*(5), 82–87. https://doi.org/10.1109/ms.2012.121

[157] OMG. (2016). *Requirements interchange format (reqif)*. Object Management Group. https://www.omg.org/reqif/

[158] SysML. (2019). *SysML Open Source Project - What is SysML? Who created it?* SysML.Org. https://sysml.org

[159] Linhares, M. V., de Oliveira, R. S., Farines, J.-M., & Vernadat, F. (2007). Introducing the modeling and verification process in SysML. *2007 IEEE Conference on Emerging Technologies & Factory Automation (EFTA 2007)*. http://dx.doi.org/10.1109/efta.2007.4416788

[160] Hause, M. (2006). The SysML Modelling Language. *Fifteenth European Systems Engineering Conference*.

[161] dos Santos Soares, M., & Vrancken, J. (2007). Requirements specification and modeling through SysML. *2007 IEEE International Conference on Systems, Man and Cybernetics*. http://dx.doi.org/10.1109/icsmc.2007.4413936

[162] Johnson, T., Kerzhner, A., Paredis, C. J. J., & Burkhart, R. (2011). Integrating models and simulations of continuous dynamics into sysml. *Journal of Computing and Information Science in Engineering*, *12*(1). https://doi.org/10.1115/1.4005452

[163] Roques, P. (2015). *How modeling can be useful to better define and trace requirements*. Modeling Requirements with SysML – Requirements Engineering Magazine. https://re-magazine.ireb.org/articles/modeling-requirements-with-sysml

[164] PMI. (2017b). *Agile practice guide*. Project Management Institute.

[165] Toussaint, M., Krima, S., Feeney, A. B., & Panetto, H. (2021). Requirement elicitation for adaptive standards development. *IFAC-PapersOnLine*, *54*(1), 863–868. https://doi.org/10.1016/j.ifacol.2021.08.101

[166] Kemmerer, R. A. (2003). Cybersecurity. 25th International Conference on Software

[167] Meisner, M. (2018). Financial consequences of cyber attacks leading to data breaches in healthcare sector. Copernican Journal of Finance & Accounting, 6(3), 63–73. https://doi.org/10.12775/cjfa.2017.017

[168] The White House. (n.d.). *Foreign Policy Cyber Security Executive Order 13636*. The White House. Retrieved May 2, 2022, from https://obamawhitehouse.archives.gov/issues/foreign-policy/cybersecurity/eo-13636

[169] The White House. (2015). *FACT SHEET: Cyber threat intelligence integration center*. Whitehouse.Gov. https://obamawhitehouse.archives.gov/the-press-office/2015/02/25/fact-sheet-cyber-threat-intelligence-integration-center

[170] Craigen, D., Diakun-Thibault, N., & Purse, R. (2014). Defining cybersecurity. *Technology Innovation Management Review*, *4*(10), 13–21. https://doi.org/10.22215/timreview/835

[171] Yampolskiy, M., Gatlin, J., & Yung, M. (2021). Myths and misconceptions in additive manufacturing security. *Proceedings of the 2021 Workshop on Additive Manufacturing (3D Printing) Security*. http://dx.doi.org/10.1145/3462223.3485618

[172] Reed, T. S. (2019). Cybercrime and Technology Losses: Claims and Potential Insurance Coverage for Modern Cyber Risks. *Tort Trial & Insurance Practice Law Journal*, *54*(1), 153–209.

[173] Tunggal, A. T. (2022). *Why is Cybersecurity Important?* UpGuard. https://www.upguard.com/blog/cybersecurity-important

[174] Gartner. (2021). *Gartner forecasts worldwide security and risk management spending to exceed $150 billion in 2021*. Gartner. https://www.gartner.com/en/newsroom/press-releases/2021-05-17-gartner-forecasts-worldwide-security-and-risk-managem

[175] McMillan, R., & Knutson, R. (2017). *Yahoo triples estimate of breached accounts to 3 billion*. The Wall Street Journal. https://www.wsj.com/articles/yahoo-triples-estimate-of-breached-accounts-to-3-billion-1507062804

[176] Fung, B. (2018). Equifax's massive 2017 data breach keeps getting worse. The Washington Post. https://www.washingtonpost.com/news/the-switch/wp/2018/03/01/equifax-keeps-finding-millions-more-people-who-were-affected-by-its-massive-data-breach/

[177] Malwarebytes. (2020). *Enduring from home - COVID-19's impact on business security*. Malwarebytes. https://www.malwarebytes.com/resources/files/2020/08/malwarebytes_enduringfromhome_report_final.pdf

[178] Verizon. (2020). Verizon: Data breach investigations report 2020. *Computer Fraud & Security*, *2020*(6), 4–4. https://doi.org/10.1016/s1361-3723(20)30059-2

[179] SBA. (n.d.). *Stay safe from cybersecurity threats*. US Small Business Administration. Retrieved May 5, 2022, from https://www.sba.gov/business-guide/manage-your-business/stay-safe-cybersecurity-threats

[180] BlueVoyant. (2021). *BlueVoyant research reveals rise in supply chain cybersecurity breaches as firms struggle to effectively monitor third-party cyber risk*. BlueVoyant. https://www.bluevoyant.com/news/bluevoyant-research-reveals-rise-in-supply-chain-cybersecurity-breaches-as-firms-struggle-to-effectively-monitor-third-party-cyber-risk/

[181] Jibilian, I., & Canales, K. (2021). The US is readying sanctions against Russia over the SolarWinds cyber attack. Here's a simple explanation of how the massive hack happened and why it's such a big deal. Insider. https://www.businessinsider.com/solarwinds-hack-explained-government-agencies-cyber-security-2020-12

[182] Make UK, AIG, & RUSI. (2018). *Cyber Security for Manufacturing*. EEF.

[183] Lezzi, M., Lazoi, M., & Corallo, A. (2018). Cybersecurity for Industry 4.0 in the current literature: A reference framework. *Computers in Industry*, *103*, 97–110. https://doi.org/10.1016/j.compind.2018.09.004

[184] Krima, S., Toussaint, M., & Feeney, A. B. (2020). Toward model-based integration specifications to secure the extended enterprise. *ASTM Smart and Sustainable Manufacturing Systems*, *4*(1), 1-1. https://doi.org/10.1520/ssms20200022

[185] IBM. (2020). *IBM: Cost of a Data Breach Report*.

[186] Kim, F. (2019). How to make sense of cybersecurity frameworks [Video]. In *RSA Conference*. https://www.youtube.com/watch?v=dt2IqidgpS4

[187] Climer, S., & Khan, M. (2019). *Cybersecurity vs regulation: Understanding dueling compliance frameworks*. Mindsight. https://gomindsight.com/insights/blog/cybersecurity-vs-regulation-compliance-dueling-frameworks/

[188] Omnistruct Marketing. (2020). *Comparing comprehensive cybersecurity frameworks.* Omnistruct. https://omnistruct.com/comparing-cybersecurity-frameworks/

[189] ISO. (2013). *ISO/IEC 27001 — Information security management*. ISO. https://www.iso.org/isoiec-27001-information-security.html

[190] 27001 Academy. (n.d.). *What is ISO 27001? A beginner's guide.* Advisera. Retrieved April 6, 2022, from https://advisera.com/27001academy/what-is-iso-27001/

[191] IT governance. (n.d.). *ISO 27001: The International Information Security Standard*. IT Governance. Retrieved April 6, 2022, from https://www.itgovernance.co.uk/iso27001

[192] ISA. (2021). New ISA/IEC 62443 standard specifies security capabilities for control system components. *International Society of Automation*. https://www.isa.org/intech-home/2018/september-october/departments/new-standard-specifies-security-capabilities-for-c

[193] ISA Global Cybersecurity Alliance. (2020). *Quick Start Guide: An Overview of the ISA/IEC 62443 Standards*. https://gca.isa.org/hubfs/ISAGCA%20Quick%20Start%20Guide%20FINAL.pdf

[194] NIST. (2018). *Getting started*. NIST. https://www.nist.gov/cyberframework/getting-started

[195] Barrett, M. P. (2018). *Framework for improving critical infrastructure cybersecurity, version 1.1*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.cswp.04162018

[196] Gordon, L. A., Loeb, M. P., & Zhou, L. (2020). Integrating cost–benefit analysis into the NIST Cybersecurity Framework via the Gordon–Loeb Model. *Journal of Cybersecurity*, *6*(1). https://doi.org/10.1093/cybsec/tyaa005

[197] NIST. (2016). *Security content automation protocol*. NIST. https://csrc.nist.gov/projects/Security-Content-Automation-Protocol

[198] OpenSCAP. (n.d.). *SCAP components*. OpenSCAP. Retrieved April 6, 2022, from https://www.open-scap.org/features/scap-components/

[199] NIST. (2016). *Security Content Automation Protocol - SCAP Specifications*. NIST. https://csrc.nist.gov/projects/security-content-automation-protocol/specifications

[200] FIRST. (n.d.). *Common vulnerability scoring system SIG*. FIRST — Forum of Incident Response and Security Teams. Retrieved April 6, 2022, from https://www.first.org/cvss/

[201] Banghart, J., Cook, M., Quinn, S., Waltermire, D., & Bove, A. (2013). *Security content automation protocol (SCAP) version 1.2 Validation program test requirements*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.ir.7511

[202] Waltermire, D. (2018). *Software identification (SWID) tagging*. NIST. https://csrc.nist.gov/projects/Software-Identification-SWID

[203] Cawthra, J., Ekstrom, M., Lusty, L., Sexton, J., & Sweetnam, J. (2020). *Data integrity: Identifying and protecting assets against ransomware and other destructive events*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.sp.1800-25

[204] Cawthra, J., Ekstrom, M., Lusty, L., Sexton, J., & Sweetnam, J. (2020). *Data integrity: Detecting and responding to ransomware and other destructive events*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.sp.1800-26

[205] Waltermire, D., Quinn, S., Scarfone, K., & Halbardier, A. (2011). *The technical specification for the Security Content Automation Protocol (SCAP) :* National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.sp.800-126r2

[206] NIST. (2018). *Uses and benefits of the framework*. NIST. https://www.nist.gov/cyberframework/online-learning/uses-and-benefits-framework

[207] NIST. (2017). *Framework for Improving Critical Infrastructure Cybersecurity*. https://csrc.nist.gov/CSRC/media/Presentations/Cybersecurity-Framework-Overview/images-media/NIST%20CSF%20Overview.pdf

[208] Quinn, S. D., Mell, P., & Kent, K. (2006). *The Security Content Automation Program (SCAP): Automating Compliance Checking, Vulnerability Management, and Security Measurement*. NIST.

[209] Filkins, B., & Wylie, D. (2019). *SANS 2019 State of OT/ICS Cybersecurity Survey*. SANS. https://www.forescout.com/resources/2019-sans-state-of-ot-ics-cybersecurity-survey/

[210] Koza, E. (2022). Semantic Analysis of ISO/IEC 27000 Standard Series and NIST Cybersecurity Framework to Outline Differences and Consistencies in the Context of Operational and Strategic Information Security. *Medicon Engineering Themes*, *2*(3).

[211] Roy, P. P. (2020, February). A high-level comparison between the NIST cyber security framework and the ISO 27001 Information Security Standard. *2020 National Conference on Emerging Trends on Sustainable Technology and Engineering Applications (NCETSTEA)*. http://dx.doi.org/10.1109/ncetstea48365.2020.9119914

[212] Mietala, A. (2020). *When should an organisation start vulnerability management?* [Master's Thesis]. https://www.theseus.fi/bitstream/handle/10024/343389/Mietala_Anssi.pdf?sequence=2

[213] Rupra, S. S., & Omamo, A. (2019). A FRAMEWORK FOR IMPROVING SECURITY IN CLOUD COMPUTING FOR SMES (FISCCS) USING SECURITY INDEX. *Global Journal of Engineering Science and Research Management*, *6*(5).

[214] Ibrahim, A., Valli, C., McAteer, I., & Chaudhry, J. (2018). A security review of local government using NIST CSF: A case study. *The Journal of Supercomputing*, *74*(10), 5171–5186. https://doi.org/10.1007/s11227-018-2479-2

[215] The White House. (2013). Executive order -- improving critical infrastructure cybersecurity. The White House. https://obamawhitehouse.archives.gov/the-press-office/2013/02/12/executive-order-improving-critical-infrastructure-cybersecurity

[216] ETSI. (2018). *TR 103 305-4 V2.1.1 - CYBER: Critical security controls for effective cyber defence - part 4: Facilitation mechanisms*. https://www.etsi.org/deliver/etsi_tr/103300_103399/10330504/02.01.01_60/tr_10330504v020101p.pdf

[217] Pylon Editorial Team. (2020). *NIST Framework*. Pylon Technology. https://pylontechnology.com/nist-framework

[218] Ferrillo, P., & Conkle, T. (2014). *Understanding and implementing the NIST cybersecurity framework*. The Harvard Law School Forum on Corporate Governance. https://corpgov.law.harvard.edu/2014/08/25/understanding-and-implementing-the-nist-cybersecurity-framework/

[219] Yehezkel, A., & Michael, T. (2014). *Cybersecurity: Breaching The Boardroom*. The Metropolitan Corporate Counsel. http://www.metrocorpcounsel.com/articles/27983/cybersecurity-breaching-boardroom/

[220] Leon, R. J. (2018). *An Event Management Framework to Aid Solution Providers in Cybersecurity* [The George Washington University]. http://pqdtopen.proquest.com/#viewpdf?dispub=10745141

[221] Berger, K. M., & Schneck, P. A. (2019). National and transnational security implications of asymmetric access to and use of biological data. *Frontiers in Bioengineering and Biotechnology*, *7*(21). https://doi.org/10.3389/fbioe.2019.00021

[222] ConnectWise. (2019). *Cybersecurity frameworks around the world*. MSSP Alert. https://www.msspalert.com/cybersecurity-news/cybersecurity-frameworks-around-the-world/

[223] Bokan, B., & Santos, J. (2021). Managing cybersecurity risk using threat based methodology for evaluation of cybersecurity architectures. *2021 Systems and Information Engineering Design Symposium (SIEDS)*, 1–6. http://dx.doi.org/10.1109/sieds52267.2021.9483736

[224] Bresnahan, E. (n.d.). *What are the benefits of the NIST cybersecurity framework*. CyberSaint Security. Retrieved April 5, 2022, from https://www.cybersaint.io/blog/benefits-of-nist-cybersecurity-framework

[225] RSI Security. (2018). How to improve your security with NIST. *RSI Security*. https://blog.rsisecurity.com/how-to-improve-your-security-with-nist/

[226] Shen, L. (2014). *The nist cybersecurity framework: Overview and potential impacts*. Aspen Publishers, Inc.

[227] Bakare, A. A. (2020). *A Methodology for Cyberthreat Ranking: Incorporating the NIST Cybersecurity Framework into FAIR Model*. University of Cincinnati.

[228] IBM Cloud Education. (2020). *NIST cybersecurity framework*. IBM. https://www.ibm.com/cloud/learn/nist-cybersecurity-framework

[229] Zaras, D. (2018). *Information Security Frameworks and Controls Catalogs*. impactmakers. http://c4jwa1cvv5kfckfu2p9gjock-wpengine.netdna-ssl.com/wp-content/uploads/2018/09/Impact-Makers_Security-Whitepaper_2018.pdf

[230] Ngoc Thach, N., Thanh Hanh, H., Ngoc Huy, D. T., Gwozdziewicz, S., Viet Nga, L. T., & Thanh Huong, L. T. (2021). TECHNOLOGY QUALITY MANAGEMENT OF THE INDUSTRY 4.0 AND CYBERSECURITY RISK MANAGEMENT ON CURRENT BANKING ACTIVITIES IN EMERGING MARKETS - THE CASE IN VIETNAM. *International Journal for Quality Research*, *15*(3), 845–856. https://doi.org/10.24874/ijqr15.03-10

[231] Dedeke, A. (2017). Cybersecurity framework adoption: Using capability levels for implementation tiers and profiles. *IEEE Security & Privacy*, *15*(5), 47–54. https://doi.org/10.1109/msp.2017.3681063

[232] HIMSS. (2018). *2018 HIMSS Cybersecurity Survey*. https://www.himss.org/sites/hde/files/d7/u132196/2018_HIMSS_Cybersecurity_Survey_Final_Report.pdf

[233] NIST. (n.d.). *Cybersecurity framework*. NIST. Retrieved April 6, 2022, from https://www.nist.gov/industry-impacts/cybersecurity-framework

[234] NIST. (2018). *Success stories*. NIST. https://www.nist.gov/cyberframework/success-stories

[235] Fukuda, Y., Kawamura, I., Kubota, Y., & Wataguchi, Y. (2019). Supply chain security measures using outcome-based approach. *Fujitsu Scientific & Technical Journal*, *55*(5), 23–29.

[236] Almagro, L. (2019). *NIST Cybersecurity Framework (CSF): A comprehensive approach to cybersecurity*. Organization of American States. https://www.oas.org/en/sms/cicte/docs/OAS-AWS-NIST-Cybersecurity-Framework(CSF)-ENG.pdf

[237] NIST. (2018). *International resources*. NIST. https://www.nist.gov/cyberframework/international-resources

[238] Mahn, A. (2019). *Picking up the framework's pace internationally*. NIST. https://www.nist.gov/blogs/cybersecurity-insights/picking-frameworks-pace-internationally

[239] NIST. (2018). *Success story: Japanese cross-sector forum*. NIST. https://www.nist.gov/cyberframework/success-stories/japanese-cross-sector-forum

[240] Matsubara, M. (2020). Cybersecurity Framework: Addressing Japan's Manpower Crunch. RSIS. https://www.think-asia.org/bitstream/handle/11540/11757/CO20018.pdf?sequence=1

[241] Brumfield, C. (2018). Why NIST is so popular in Japan. *CyberScoop*. https://www.cyberscoop.com/nist-japan-workforce/

[242] CIS Sapienza. (2016). *2015 Italian Cyber Security Report*. CIS Sapienza. https://www.cybersecurityframework.it/sites/default/files/CSR2015_ENG.pdf

[243] Ciglic, K. & Microsoft. (n.d.). *Cybersecurity policy framework*. Microsoft Cybersecurity. Retrieved April 7, 2022, from https://www.microsoft.com/en-us/cybersecurity/content-hub/cybersecurity-policy-framework

[244] Angelini, M., Ciccotelli, C., Franchina, L., Marchetti-Spaccamela, A., & Querzoni, L. (2020). Italian National Framework for Cybersecurity and Data Protection. In *Privacy Technologies and Policy* (pp. 127–142). Springer International Publishing. http://dx.doi.org/10.1007/978-3-030-55196-4_8

[245] Nieves, M. (2019). Uruguay y la ciberseguridad: entre los determinismos regionales y el proceso doméstico. *IX Encuentro Del CERPI y La VII Jornada Del CENSUD*. http://sedici.unlp.edu.ar/bitstream/handle/10915/116258/Documento_completo.pdf-PDFA.pdf?sequence=1

[246] AGESIC. (2020). *Marco de Ciberseguridad*. Agencia de Gobierno Electrónico y Sociedad de La Información y Del Conocimiento. https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/comunicacion/publicaciones/marco-ciberseguridad

[247] ASIC. (2015). *Report REP 429 - Cyber resilience: Health check*. Australian Securities and Investments Commission. https://asic.gov.au/media/3062900/rep429-published-19-march-2015-1.pdf

[248] ASIC. (2019). *Report REP 651 - Cyber resilience of firms in Australia's financial markets: 2018–19*. https://download.asic.gov.au/media/5416529/rep651-published-18-december-2019.pdf

[249] Benz, M., & Chatterjee, D. (2020). Calculated risk? A cybersecurity evaluation tool for SMEs. *Business Horizons*, *63*(4), 531–540. https://doi.org/10.1016/j.bushor.2020.03.010

[250] Santos, J. (2020). A beginners guide to the NIST cybersecurity framework. *MapleTronics* . https://www.mapletronics.com/post/a-beginners-guide-to-the-nist-cybersecurity-framework

[251] Angelini, M., Lenti, S., & Santucci, G. (2017). CRUMBS: A cyber security framework browser. *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*. http://dx.doi.org/10.1109/vizsec.2017.8062194

[252] Teodoro, N., Goncalves, L., & Serrao, C. (2015). NIST cybersecurity framework compliance: A generic model for dynamic assessment and predictive requirements. *2015 IEEE Trustcom/BigDataSE/ISPA*. http://dx.doi.org/10.1109/trustcom.2015.402

[253] Abraham, C., Chatterjee, D., & Sims, R. R. (2019). Muddling through cybersecurity: Insights from the U.S. healthcare industry. *Business Horizons*, *62*(4), 539–548. https://doi.org/10.1016/j.bushor.2019.03.010

[254] Dimensional Research. (2016). NIST cybersecurity framework adoption linked to higher security confidence according to new research from tenable network security. In *Tenable®*. Tenable. https://static.tenable.com/marketing/tenable-csf-report.pdf

[255] Schlimmer, S. (2017). *Five steps to simplify NIST cybersecurity framework adoption*. Infosecurity Magazine. https://www.infosecurity-magazine.com/opinions/simplify-nist-cybersecurity/

[256] Subedi, H. (2021). What is the NIST cybersecurity framework and how to get started. *Jones IT*. https://www.itjones.com/blogs/2021/11/1/what-is-the-nist-cybersecurity-framework-and-how-to-get-started

[257] Roy, P. P. (2020). A high-level comparison between the NIST cyber security framework and the ISO 27001 Information Security Standard. *2020 National Conference on Emerging Trends on Sustainable Technology and Engineering Applications (NCETSTEA)*. http://dx.doi.org/10.1109/ncetstea48365.2020.9119914

[258] Harry, C. T., & Gallagher, N. W. (2019). *An Effects-Centric Approach to Assessing Cybersecurity Risk*. CISSM Report. https://spp.umd.edu/sites/default/files/2019-07/cissm_risk_framework_overview_final_030119v2.pdf

[259] Priyadarshini, I., Kumar, R., Tuan, L. M., Son, L. H., Long, H. V., Sharma, R., & Rai, S. (2021). A new enhanced cyber security framework for medical cyber physical systems. *SICS Software-Intensive Cyber-Physical Systems*, *35*(3–4), 159–183. https://doi.org/10.1007/s00450-021-00427-3

[260] Lee, I. (2021). Cybersecurity: Risk management framework and investment cost analysis. *Business Horizons*, *64*(5), 659–671. https://doi.org/10.1016/j.bushor.2021.02.022

[261] Stouffer, K., Zimmerman, T., Tang, C., Lubell, J., Cichonski, J., & McCarthy, J. (2017). *Cybersecurity framework manufacturing profile*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.ir.8183

[262] Barker, W. C., Fisher, W., Scarfone, K., & Souppaya, M. (2022). *Ransomware risk management: A cybersecurity framework profile*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.ir.8374

[263] Marron, J., Gopstein, A., Bartol, N., & Feldman, V. (2019). *Cybersecurity framework smart grid profile*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.tn.2051

[264] IBM. (2014). *IBM Security Services 2014 cyber security intelligence index*. IBM. https://i.crn.com/sites/default/files/ckfinderimages/userfiles/images/crn/custom/IBMSecurityServices2014.PDF

APPENDIX – SUMMARY TABLE OF THE CYBERSECURITY FRAMEWORKS REVIEW

Frameworks	Criterion (1)	Criterion (2)	Criterion (3)	Criterion (4)
American National Standards Institute (ANSI) framework	+	+	+	-
Australian Signals Directorate (ASD) Essential 8	+	+	-	
Center for Internet Security (CIS) Controls	+	+	-	
Cloud Security Alliance (CSA) Cloud Controls Matrix (CCM)	+	-		
Committee of Sponsoring Organizations (COSO)	+	+	-	
Consortium for IT Software Quality (CISQ)	+	-		
Control Objectives for Information Technology (COBIT)	+	+	-	
Cybersecurity and Infrastructure Security Agency (CISA) Transportation Systems Sector (TSS) Cybersecurity Framework	+	-		
Cybersecurity Maturity Model Certification (CMMC)	+	+	+	-
European Telecommunications Standards Institute (ETSI) TC CYBER (Technical Committee on Cyber Security) framework	+	+	+	-
European Union Agency for Cybersecurity (ENISA) National Capabilities Assessment Framework	+	+	-	
Factor Analysis of Information Risk (FAIR) Cyber Risk Framework	+	+	+	-

Frameworks	Criterion (1)	Criterion (2)	Criterion (3)	Criterion (4)
Federal Risk and Authorization Management Program (FedRAMP)	+	-		
Federal Information Systems Management Act (FISMA)	-			
General Data Protection Regulation (GDPR)	-			
Health Insurance Portability and Accountability Act (HIPAA)	-			
HITRUST Cybersecurity Framework (CSF)	+	-		
International Association of Medical Science Educators (IASME) Governance	+	+	+	-
Information Security Forum (ISF) Standard of Good Practice for Information Security (SOGP 2020)	+	+	+	-
International Organization for Standardization (ISO) 27001	+	+	+	+
International Society of Automation (ISA/IEC 62443)	+	+	+	+
International Telecommunications Union (ITU) National Cybersecurity/ Critical Information Infrastructure Protection (CIIP)	+	+	-	
Internet of Things (IoT) Cybersecurity Alliance (IOTCA)	+	-		
Internet of Things (IoT) Security Foundation (IoTSF) Security Compliance Framework	+	-		
MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK)	+	+	-	
National Cyber Security Centre (NCSC) Cyber Assessment Framework	+	+	-	

Frameworks	Criterion (1)	Criterion (2)	Criterion (3)	Criterion (4)
(CAF)				
National Institute of Standards and Technology (NIST) Cybersecurity Framework (CSF)	+	+	+	+
New Zealand Protective Security Requirements (PSR)	+	+	-	
NIST 800-53 Cybersecurity Framework	+	+	+	-
North American Electric Reliability Corporation (NERC)	+	-		
New York Department of Financial Services (NY DFS)	-			
Payment Card Industry Data Security Standard (PCI DSS)	-			
Saudi Arabian Monetary Authority (SAMA) Cybersecurity Framework	+	-		
Security Content Automation Protocol (SCAP)	+	+	+	+
Service Organization Control (SOC) 2	+	+	-	
Ten Steps to Cybersecurity	+	+	-	
36	**31**	**22**	**11**	**4**

+ = meet the criterion; - = didn't meet the criterion

gray cell = didn't meet one of the previous criteria

blue line = final set

RÉSUMÉ ÉTENDU EN FRANÇAIS

Au fil des siècles, les progrès technologiques ont modifié les méthodes de production utilisées par les humains. Les nouvelles techniques et processus de production ont radicalement changé les conditions de travail et les modes de vie des gens.

La première révolution industrielle a marqué la naissance de la mécanisation grâce à l'utilisation de l'eau et de la vapeur. La deuxième révolution industrielle a reflété l'émergence d'une production de masse possible grâce à la découverte de l'électricité. La troisième a marqué l'émergence de l'automatisation des processus de production grâce à l'introduction de l'électronique et des technologies de l'information. Enfin, la quatrième révolution industrielle, également connue sous le nom d' « Industrie 4.0 », a été façonnée par la révolution numérique basée sur les systèmes cyber-physiques (CPS). Elle est également caractérisée par l'interconnectivité des systèmes et l'accès aux données en temps réel.

La révolution numérique dans le monde de l'industrie est alimentée par les progrès des technologies de l'information et de la communication. Sur le même principe, les procédés automatisés de collecte et d'analyse des données en temps réel ont succédé aux méthodes manuelles autrefois utilisées.

La numérisation de l'industrie, l'adoption de systèmes cyber-physiques et de la technologie IoT/CPS ont facilité et entraîné la génération de grands volumes de données hétérogènes. Les organisations produisent, consomment et échangent des volumes massifs de données dans le cadre de leurs opérations quotidiennes. Les données ont désormais le pouvoir de se transformer instantanément en informations, en connaissances et en décisions raisonnées, dans le but d'améliorer les performances des organisations (par exemple, réduire les coûts ou optimiser les ressources). Les données numériques sont devenues un acteur essentiel dans de nombreux processus décisionnels et un catalyseur essentiel pour améliorer la compétitivité de l'industrie.

L'industrie 4.0 (ou la soi-disant Industrie du Futur en France) nécessite et permet un accès et un échange rapides de ces données entre une variété d'applications et de systèmes d'information - au sein et entre les organisations. Un produit crée et dépend d'une grande quantité de données au cours de son cycle de vie en réponse à différents processus (par exemple, conception, fabrication, inspection) et besoins (par exemple, techniques, commerciaux, réglementaires). Chaque organisation impliquée dans le cycle de vie du produit s'appuie sur ces données pour remplir ses fonctions. Elles représentent le « carburant » derrière la contribution, l'efficacité et la valeur des organisations : les organisations peuvent ainsi créer plus de valeur et accélérer l'innovation en échangeant des données entre elles, ce qui facilite la collaboration.

La transformation numérique de l'industrie a conduit à des environnements et des processus plus connectés, automatisés et centrés sur les données. Les données sont devenues un catalyseur clé des processus, des échanges et de la prise de décision. L'industrie dépend fortement des données et de l'échange de ces données entre les différentes machines, systèmes et parties prenantes. L'échange de données, par définition, est le processus d'envoi et de réception de données de manière à ce que le contenu ou la signification des données n'ait pas été altéré lors de la transmission, en d'autres termes que les données reçues soient une représentation exacte des données envoyées.

La numérisation de l'industrie a souligné l'importance de la gestion de l'information, de l'échange de données et de l'interopérabilité des différents acteurs dans les processus de fabrication. L'émergence de nouvelles technologies et de sources de données en réseau offre de nouvelles possibilités de collaboration organisationnelle grâce à l'échange de données à grande vitesse et à grand volume. En d'autres termes, cette ère numérique a contribué à améliorer la vitesse, le volume, la précision et la cohérence des échanges de données et des innovations entre et au sein des organisations. Les données sont désormais disponibles et partageables à un rythme plus rapide que les organisations ne peuvent les consommer et les suivre.

Malheureusement, un échange de données rapide et fiable est aussi une opération complexe qui s'accompagne de multiples défis, chacun pouvant avoir des conséquences drastiques sur les organisations, leurs opérations, leurs produits et leurs collaborateurs. Dans cette thèse, nous nous concentrerons sur les menaces croissantes accompagnant les changements rapides de l'environnement dynamique de l'industrie, la sécurité des échanges de données, et démontrerons comment une meilleure interopérabilité et traçabilité des données peut y faire face. L'interopérabilité des données et la traçabilité des données sont deux défis majeurs inhérents au volume massif d'échanges de données induit par la numérisation industrielle de l'Industrie 4.0, auxquels les industriels sont confrontés et doivent comprendre pour y faire face.

Interopérabilité des données

Suite à cette transformation numérique de l'industrie et à la modernisation des technologies de communication adoptées, les données sont désormais disponibles de tous, à tous, et dans une multitude de formats. Les organisations peuvent facilement connecter différents logiciels et systèmes physiques, en interne et au sein de leur réseau de collaborateurs, tant que ces systèmes parlent un langage commun.

Malheureusement, les organisations de l'industrie d'aujourd'hui sont caractérisées par des environnements complexes constitués de composants spécifiques à un domaine tels que des systèmes, des réseaux ou des machines, regroupés en groupes hétérogènes. Bien que l'interaction de ces composants soit cruciale pour l'industrie car elle soutient les processus de production, une interopérabilité efficace entre

tous les éléments du cycle de vie du produit est un défi croissant. La quantité de données produites et consommées continue d'augmenter en raison de cet écosystème croissant (de machines, de systèmes et de réseaux), mais aussi du nombre de formats de données. Ces données sont collectées à partir de sources de données distribuées et ne partagent donc pas nécessairement le même format. L'hétérogénéité des données est un facteur important dans l'échange de données. Les différents composants de l'environnement d'une organisation doivent être capables d'interpréter, d'utiliser, d'intégrer et de comparer sans ambiguïté les informations échangées.

Ces différents systèmes ont besoin d'un langage commun pour échanger et comprendre les informations. L'utilisation de normes de données neutres basées sur des modèles aide à fournir un format de données commun et facilite ainsi l'interopérabilité technique entre toutes les parties impliquées dans un échange. Les normes d'information sont essentielles pour intégrer, échanger et interpréter correctement les données sur lesquelles les industriels s'appuient. Les normes définissent un langage convenu (format de données, définitions, etc.) pour l'échange de données entre les différents systèmes qui consomment, traitent et produisent des données. Le manque de standardisation se traduit par une multiplication des formats d'information qui ne sont pas forcément compatibles entre eux, rendant difficile la communication et l'échange de données pour les acteurs.

Les normes d'information sont un atout important pour les organisations car elles facilitent l'interaction commerciale et la compatibilité entre les systèmes, ainsi que l'interopérabilité entre les systèmes, les personnes et les organisations. La standardisation des informations permet également de gagner du temps et de réduire les coûts en éliminant le besoin d'avoir des traducteurs séparés pour chaque paire de systèmes devant échanger des données. L'adoption et la mise en œuvre de normes par les organisations améliorent leur performance, compétitivité et transparence, étant donné que les normes favorisent l'accessibilité de l'information par toutes les parties prenantes. Les normes d'information sont de puissants outils de productivité et d'innovation car elles fournissent un langage, des vocabulaires et des plateformes communs permettant aux organisations de collaborer et de développer de nouvelles solutions technologiques. Les normes soutiennent également l'innovation en augmentant la vitesse à laquelle les organisations mettent de nouveaux produits sur le marché, grâce à une collaboration plus rapide et plus précise. Les normes d'information sont donc des catalyseurs clés de l'évolution et de la numérisation du secteur de l’industrie.

Malgré cela, les normes d'information présentent un défi majeur, qui peut avoir un impact sur leur adoption et leur mise en œuvre par les organisations : la complexité et la durée actuelle du processus de développement de normes d’information majeures sont incompatibles non seulement avec les besoins et le rythme de l'industrie, mais aussi avec la durée de vie des données.

Le processus de développement des normes d'information est complexe. Ce processus est généralement long, irrégulier et difficile à planifier. Premièrement, la méthodologie *waterfall* pour la gestion de projet est prédominante, ce qui implique que 1) l'intégralité du livrable n'est disponible (pour révision) qu'à la fin de l'itération de développement, et 2) les exigences doivent être définies au début du projet et ne changent pas pendant toute l'itération. Selon ISO, les itérations de développement des normes peuvent durer entre 18 et 48 mois. Cela signifie qu'une nouvelle exigence identifiée après le début d'une nouvelle itération ne sera pas traitée avant 18 à 48 mois. Mais dans les deux cas, un délai supplémentaire doit être accordé aux éditeurs de logiciels pour mettre en œuvre, tester et fournir des solutions logicielles mises à jour.

Deuxièmement, les normes sont élaborées par des experts qui travaillent pour différentes organisations. La contribution et la participation de ces experts au processus d'élaboration des normes sont entièrement volontaires. Les ressources disponibles dépendent des emplois du temps des experts et des besoins de leurs organisations, ce qui rend le processus de développement irrégulier et difficile à planifier.

La durée et la gestion du processus d'élaboration des normes ne correspondent pas aux besoins de l'industrie. La forte concurrence du marché se traduit par des cycles de vie des produits raccourcis et des exigences qui changent souvent et plus rapidement que le rythme de développement des normes. Le processus de développement des normes est aussi incompatible avec la durée de vie des données. L'industrie 4.0 valorise la vitesse et l'innovation rapide. Par conséquent, les industriels ont besoin que les organismes de développement de normes accélèrent et simplifient le processus développement des normes, de sorte que les normes qui en résultent représentent les besoins actuels de l'industrie et soient plus rapidement adoptées.

Pour résumer, l'innovation et la collaboration rapides sont entravées par les défis d'interopérabilité des données. Une collaboration accrue s'accompagne d'un volume croissant d'échanges de données qui est associé à un nombre plus grand de systèmes hétérogènes qui doivent communiquer ensemble et se comprendre. Les normes d'information sont une solution reconnue pour soutenir l'interopérabilité entre les systèmes et le partage significatif des données, car elles fournissent un langage commun aux systèmes d'information qui prennent en charge les activités et les opérations des organisations. Le processus de développement des normes est long et complexe, ce qui empêche les normes de suivre le rythme rapide de l'environnement qu'elles ont besoin de supporter et pour lesquels elles ont besoin d'assurer l'interopérabilité, ainsi que le rythme des innovations technologiques et des cybermenaces en constante évolution. Le processus de développement des normes est confronté à des inefficacités qui ralentissent la publication et l'adoption des normes.

Dans cette thèse, nous proposons une transition de la gestion de projet prédictive à adaptative avec l'utilisation des méthodes Agile. Nous présentons quelques pratiques Agile, à travers un cas de test utilisant

STEP AP242, qui peuvent aider les équipes de développement standard à surmonter certaines de ces inefficacités. Agile fournit un moyen (par le biais de processus et d'outils) pour 1) améliorer la gestion des ressources disponibles (c'est-à-dire le personnel bénévole) en réduisant le besoin de planification et d'engagement à long terme, et 2) améliorer la collaboration entre ces communautés en offrant un canal de communication plus rapide. Malgré ces avantages, Agile s'accompagne également de défis. Nous avons identifié un besoin clair pour une meilleure traçabilité et visibilité des exigences, de leur élicitation, de leur gestion et de leur mise en œuvre. Pour répondre à ce besoin, nous avons développé et proposé un nouveau modèle d'élicitation des exigences qui sert de base pour fournir une traçabilité et une visibilité des exigences et des décisions, en enregistrant et en exploitant les comptes-rendus de réunions de projet de manière formelle. Nous avons également proposé une implémentation de ce modèle via une personnalisation de Jira, un outil utilisé par les équipes de développement STEP pour le suivi des bugs et des problèmes.

Traçabilité des données

L'industrie est devenue plus automatisée, connectée et centrée sur les données. La mise en réseau des machines, des systèmes et des produits, et l'émergence d'environnements cyber-physiques permettent aux données d'être plus rapidement accessibles et facilitent l'échange rapide et opportun de données entre les systèmes qui nécessitent l'information et les systèmes qui disposent de l'information. Ces échanges de données sont à la fois intra- et inter-organisationnels et peuvent être caractérisés comme des échanges à haut débit, à volume élevé, à haute fréquence et à faible latence. Les données sont désormais utilisées pour générer et partager des décisions à un rythme et à un volume nettement supérieurs à tout ce que les humains peuvent manuellement valider ou suivre.

Ces rythme et volume d'échange de données dans le monde de l'industrie s'accompagnent de défis importants. La forte dépendance à l'égard des décisions basées sur les données, et l'intégration de nouvelles technologies ont rendu les organisations plus vulnérables aux cybermenaces, une préoccupation majeure pour les entreprises, quelles que soient leur taille et leur secteur. Le secteur manufacturier génère de grandes quantités de données et en dépend fortement, ce qui fait de ce secteur une cible idéale pour les cyberattaques. Sans surprise, le secteur manufacturier a été particulièrement touché par la cybercriminalité en 2020 et 2021. Selon le rapport X-ForceThreat Intelligence Index 2022 d'IBM Security, le secteur manufacturier a été le secteur le plus attaqué en 2021 (avec 23,2 % de toutes les attaques) alors qu'il était classé deuxième en 2020 (avec 17,7 % de toutes les attaques) et huitième en 2019 (avec 8,1 % de toutes les attaques).

Généralement, les menaces de sécurité sont classées selon les principes directeurs du modèle de sécurité de la triade CIA : confidentialité, intégrité et disponibilité *(confidentiality, integrity and*

availability). La confidentialité des données exige que les données restent secrètes ou privées, l'intégrité des données exige que les données soient fiables et exemptes de falsification, et enfin, la disponibilité des données exige que les données soient toujours accessibles à un accès autorisé en cas de besoin. Les menaces et la vulnérabilité sont évaluées en fonction du type de risques associés et des dommages potentiels qu'ils peuvent causer aux actifs d'une organisation tels que les données, les applications et les systèmes.

Ces risques ne peuvent pas toujours être évités et constituent un défi important à identifier et à contenir. Le *Cost of a data breach report 2022* d'IBM montre qu'en 2021, le temps moyen d'identification (MTTI) d'une violation de données était de 212 jours et de 75 jours pour être contenue (MTTC) pour un cycle de vie total de 287 jours. Cela représente une légère augmentation par rapport à 2020, lorsque le délai moyen pour identifier et contenir une violation de données était de 280 jours (avec une moyenne de 207 jours pour identifier et une moyenne de 73 jours pour contenir).

Une menace particulièrement pertinente est la manipulation de données, une attaque qui se concentre sur la modification subtile des données dans le but de manipuler les décisions basées sur les données et qui s'appuie sur l'échange de données pour propager des données et des décisions falsifiées dans une organisation et son réseau. Cette falsification peut entraîner la corruption, la modification et/ou la destruction des données, entraînant finalement une perte de confiance dans les données et les décisions qui en découlent. La manipulation des données peut conduire à des décisions et des processus corrompus, ainsi qu'à des chaînes d'approvisionnement entières compromises, ce qui peut avoir de graves conséquences pour les organisations, telles que des dommages économiques et de réputation.

Lorsque des données sont échangées, elles quittent le système privé et sécurisé du propriétaire des données pour être envoyées à d'autres systèmes. Ce processus présente le risque critique que les données échangées aient été falsifiées par des parties non autorisées. Il est donc important de s'assurer que les données restent exactes, authentiques et dignes de confiance tout au long du processus d'échange. L'intégrité des données est l'aspect de la triade CIA le plus impacté par les échanges de données, car l'intégrité aide à la fois l'expéditeur, qui doit s'assurer que les données qui lui sont attribuées ne sont pas altérées, et le destinataire, qui a besoin « de la garantie que le message qui est envoyé est le même que le message reçu et que le message n'ait pas été altéré en transit ».

L'intégrité des données présente deux défis principaux : 1) valider l'exactitude des données et 2) détecter les inexactitudes. Le premier est généralement résolu à l'aide de signatures numériques, tandis que le second est plus complexe et doit encore être résolu. La complexité (c'est-à-dire le nombre d'acteurs et d'étapes impliquées), le rythme et le volume des échanges de données auxquels les organisations participent rendent impossible la prise en compte et le suivi manuels de chaque inexactitude. Ces mêmes bénéfices et

avantages qui rendent les fabricants plus compétitifs et innovants les rendent également plus vulnérables aux attaques contre l'intégrité des données.

Pour résumer, les défis de la traçabilité des données ont un impact sur la fiabilité des données et des décisions, un élément clé pour obtenir des avantages tels que l'optimisation des chaînes d'approvisionnement. L'exploitation des nouvelles technologies de l'information et de la communication a permis une nouvelle approche axée sur les données pour les décisions et les processus commerciaux, mais elle a également considérablement augmenté le nombre et la complexité des défis de cybersécurité associés. Les données sont désormais exposées à de nombreuses cybermenaces, y compris la manipulation de données. Les décisions basées sur les données sont vulnérables à la vitesse à laquelle les données falsifiées peuvent se propager dans les organisations pour les corrompre. La croissance constante des données produites et échangées est susceptible de faire encore augmenter le temps moyen pour identifier de telles menaces. Les organisations mettent généralement plus de temps à se remettre des événements de manipulation de données car 1) contrairement aux ransomware, ils mettent du temps à être identifiés, 2) ils utilisent l'échange de données, un élément clé de l'industrie 4.0, comme vecteur de propagation, et 3) la propagation rend difficile la déterminer l'étendue et l'impact de ces attaques. Un cadre formel de traçabilité des données est toujours nécessaire pour suivre l'échange de données à travers des réseaux d'organisations vastes et complexes afin d'identifier et de contenir la propagation de données falsifiées.

Cette thèse a analysé les cadres existants de cybersécurité et leurs limites, et a proposé un nouveau cadre, sous la forme d'un profil étendu du cadre de cybersécurité du NIST, pour atténuer, gérer et suivre les attaques de manipulation de données. Ce profil vise à aider les organisations à aligner leurs exigences de prévention et d'atténuation des risques liés à l'intégrité des données et leurs ressources avec les éléments du cadre de cybersécurité du NIST, ainsi qu'à évaluer leur état de préparation face à de telles attaques. Ce profil est accompagné de recommandations de mise en œuvre pour faciliter son adoption et sa mise en œuvre. Ces recommandations de mise en œuvre complètent le profil de développement d'une solution formelle de traçabilité des données pour détecter et atténuer les événements de manipulation des données. Ces recommandations consistent en des témoignages d'utilisateurs, des mesures de sécurité et des directives pour la formation à la sensibilisation aux flux d'informations.

Les données numériques sont un atout essentiel pour les organisations car elles soutiennent la prise de décision et la stratégie organisationnelles. Les organisations doivent gérer et échanger de gros volumes de données tout en préservant l’intégrité, la qualité et la fiabilité de leurs données. Les échanges de données doivent être précis et compréhensibles par tous les systèmes organisationnels et les parties prenantes pour que les décisions et les processus basés sur ces échanges soient fiables et dignes de confiance. Par

conséquent, dans cette thèse, nous avons abordé deux défis principaux et interdépendants liés à l'échange de données : l'interopérabilité des données et la traçabilité des données. Cette thèse se concentre sur la facilitation de l'échange de données tout en empêchant la manipulation des données grâce à l'utilisation de normes d'information.

Pour conclure, cette thèse vise à réduire le temps 1) de développement et de mise en œuvre de solutions d'interopérabilité des données et 2) des opérations de traçabilité des données en réponse aux cyber-attaques dont sont victimes les industriels, afin de répondre à la vitesse à laquelle les échanges de données sont désormais mis en place et réalisés.