UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

ERNESTO FONSECA VEIGA

SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems

Goiânia
2025

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

[ ] Dissertação [X] Tese [ ] Outro*:_____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Ernesto Fonseca Veiga

3. Título do trabalho

SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento [X] SIM [ ] NÃO[1]

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:
a) consulta ao(à) autor(a) e ao(à) orientador(a);
b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
O documento não será disponibilizado durante o período de embargo.
Casos de embargo:
- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.

Documento assinado eletronicamente por **Renato De Freitas Bulcao Neto**, **Professor do Magistério Superior**, em 19/11/2025, às 11:46, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020 .

Documento assinado eletronicamente por **Ernesto Fonseca Veiga**, **Discente**, em 24/11/2025, às 16:09, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5799716** e o código CRC **0FB5A0DB**.

Referência: Processo nº 23070.045793/2025-24 SEI nº 5799716

ERNESTO FONSECA VEIGA

SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems

Thesis presented to the Postgraduate Program in Computer Science of the Instituto de Informática of the Universidade Federal de Goiás, as a partial requirement for obtaining a Ph.D in Computer Science.

Concentration area: Computer Science.
Research line: Computing Methodologies and Techniques
Advisor: Prof. Dr. Renato de Freitas Bulcão Neto

Goiânia
2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Veiga, Ernesto Fonseca
SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems [tese] = SafeSecRETS: Suporte Integrado para Engenharia de Requisitos de Segurança e Proteção em Sistemas IoT Críticos / Ernesto Fonseca Veiga. - 2025.
CCLXXX, 280 f.: 2025

Orientador: Prof. Dr. Renato de Freitas Bulcão Neto
Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Informática (INF), Programa de Pós-Graduação em Ciência da Computação, Goiânia, 2025.
Apêndice.
Bibliografia.
Inclui: siglas, lista de figuras, lista de tabelas.

1. Engenharia de Requisitos. 2. Sistemas IOT Críticos. 3. Segurança e Proteção. 4. Artefatos. 5. Canvas e STPA.

I. Bulcão Neto, Renato de Freitas, orient. II. Título.

CDU 004

UNIVERSIDADE FEDERAL DE GOIÁS

INSTITUTO DE INFORMÁTICA

ATA DE DEFESA DE TESE

Ata nº **29** da sessão de Defesa de Tese de **Ernesto Fonseca Veiga**, que confere o título de Doutor em Ciência da Computação, na área de concentração em Ciência da Computação.

Aos vinte e um dias do mês de de dois mil e vinte e cinco, a partir das nove horas, no lab. 250, realizou-se a sessão pública de Defesa de Tese intitulada **"SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems"**. Os trabalhos foram instalados pelo Orientador, Professor Doutor Renato de Freitas Bulcão Neto (INF/UFG) com a participação dos demais membros da Banca Examinadora: Professor Doutor Iwens Gervásio Sene Junior (INF/UFG), membro titular externo; Professora Doutora Isabel Sofia Sousa Brito (Instituto Politécnico de Beja), membra titular externa; Professor Doutor Jaelson Freire Brelaz de Castro (UFPE), membro titular externo e Professor Doutor Johnny Cardoso Marques (ITA), membro titular externo. A participação dos professores Jaelson Freire Brelaz de Castro, Isabel Sofia Sousa Brito e Johnny Cardoso Marques ocorreu por meio de videoconferência. Durante a arguição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese, tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor Renato de Freitas Bulcão Neto, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos vinte e um dias do mês de de dois mil e vinte e cinco.

TÍTULO SUGERIDO PELA BANCA

Documento assinado eletronicamente por **Jaelson Feire Brelaz de Castro**, **Usuário Externo**, em 21/10/2025, às 12:14, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020 .

Documento assinado eletronicamente por **Renato De Freitas Bulcao Neto**, **Professor do Magistério Superior**, em 21/10/2025, às 12:15, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020 .

Documento assinado eletronicamente por **Ernesto Fonseca Veiga**, **Discente**, em 21/10/2025, às 12:19, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Isabel Sofia Sousa Brito**, **Usuário Externo**, em 21/10/2025, às 15:03, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Johnny Cardoso Marques**, **Usuário Externo**, em 22/10/2025, às 14:28, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020 .

Documento assinado eletronicamente por **Iwens Gervasio Sene Junior**, **Professor do Magistério Superior**, em 05/01/2026, às 17:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020 .

A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5730376** e o código CRC **15DC2622**.

Referência: Processo nº 23070.045793/2025-24 SEI nº 5730376

ERNESTO FONSECA VEIGA

SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems

Thesis defended in the Postgraduate Program in Computer Science of the Instituto de Informática of the Universidade Federal de Goiás as a partial requirement for obtaining the title of Doctor in Computer Science, approved on October 21, 2025, by the Examining Board constituted by the professors:

Prof. Dr. Renato de Freitas Bulcão Neto
Instituto de Informática
Universidade Federal de Goiás (UFG)
President of the Examining Board

Prof. Dr. Jaelson Freire Brelaz de Castro
Centro de Informática (CIn)
Universidade Federal de Pernambuco (UFPE)
External Member

Profa. Dra. Isabel Sofia Sousa Brito
Escola Superior de Tecnologia e Gestão (ESTIG)
Instituto Politécnico de Beja (IPBeja)
External Member

Prof. Dr. Johnny Cardoso Marques
Divisão de Ciência da Computação (IEC)
Instituto Tecnológico de Aeronáutica (ITA)
External Member

Prof. Dr. Iwens Gervásio Sene Júnior
Instituto de Informática (INF)
Universidade Federal de Goiás (UFG)
Internal Member

Ernesto Fonseca Veiga

He holds a Bachelor's degree in Informatics from the Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG) and a Master's degree in Computer Science from the Instituto de Informática (INF) at the Universidade Federal de Goiás (UFG). He has worked as a researcher at the Participatory Public Policy Laboratory (L3P) at UFG, as Director of Innovation at the Secretaria de Ciência, Tecnologia e Inovação (SCTI) of the Prefeitura Municipal de Aparecida de Goiânia, and as a temporary professor at INF/UFG. He is currently a professor at the Pontifícia Universidade Católica de Goiás (PUC-GO). His research interests include Software Engineering, with an emphasis on Requirements Engineering, Safety and Security, and IoT systems.

Acknowledgments

First of all, I thank God for allowing me to reach this point. Without His guidance, none of this would have been possible.

I would like to express my gratitude to my advisor, Professor Renato Bulcão, for once again giving me the opportunity to be his student and, above all, for his valuable guidance and lessons that have shaped me as a researcher. I also extend my thanks to Professor Taciana Kudo, who contributed to this mentoring process at key moments during my doctoral journey.

I sincerely thank the professors serving on the examination panel for their availability, guidance, and valuable contributions to the development of this work.

I am grateful to all members of the research group with whom I have shared these last few years for their friendship and collaboration: Thalia, Mariana, Lívia, Cinara, and Filipe. My gratitude also goes to all my friends, both from academia and beyond, who encouraged and supported me throughout this process.

I also thank the entire INF/UFG team, professors, the PPGCC secretariat, and the coordination, for their support, guidance, and willingness to assist me. I am grateful to PUC Goiás and the staff of the Polytechnic and Arts School for their trust and support.

I am deeply thankful to my parents, Raimundo and Eva, for their encouragement, unconditional support, and exemplary guidance. My gratitude extends to my siblings and all family members who supported me throughout this journey; I apologize for the many absences that this process required.

Finally, I owe my profound gratitude to my wife, Karlla Loane, who has been my partner and greatest encourager from the first day to the last. Her support, patience, and understanding were essential, giving me strength even in the most challenging moments. Thank you for sharing every dream with me, including the one of completing this journey.

Resumo

Veiga, Ernesto. **SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems**. Goiânia, 2025. 280p. Tese de Doutorado. Programa de Pós-Graduação em Ciência da Computação, Instituto de Informática, Universidade Federal de Goiás.

[**Contexto**] As características de autonomia e conectividade inerentes aos sistemas IoT críticos têm exigido o tratamento conjunto de safety e security desde as etapas iniciais de projeto, a fim de prevenir potenciais perdas. [**Problema**] Entretanto, apesar da crescente preocupação da literatura com o tratamento integrado desses requisitos, ainda são escassos os esforços de pesquisa que abordam de forma sistemática as atividades e tarefas do processo de Engenharia de Requisitos (ER) de safety e security, visando produzir a documentação necessária para orientar as demais fases do ciclo de vida de desenvolvimento. [**Objetivo**] Para preencher essa lacuna, esta pesquisa propõe artefatos de apoio ao processo de ER de safety e security em sistemas IoT críticos, contemplando as etapas de planejamento de projeto, elicitação, análise e especificação dos requisitos. A proposta realiza um processo de coanálise desses requisitos, de modo a tratar suas relações desde as fases iniciais do projeto. [**Métodos**] O trabalho foi conduzido segundo a metodologia de Design Science Research (DSR), que orientou a proposição, construção e avaliação dos artefatos, os quais incluem: (i) um modelo para o planejamento de projetos de sistemas IoT críticos (*SafeSecIoT Canvas*), instanciado a partir de um metamodelo para suporte metodológico à construção de artefatos baseados em canvas (*MM4Canvas*); (ii) um método que estende o System Theoretic Process Analysis (STPA) para a análise conjunta de safety e security (*STPA-SafeSecIoT*); e (iii) uma ferramenta que integra os artefatos (i) e (ii) em apoio às atividades e tarefas do processo de ER (*SafeSecRETS*). [**Resultados**] As avaliações realizadas indicam que os artefatos propostos apoiam de maneira eficaz a execução das atividades e tarefas do processo de ER de safety e security, apresentando alta utilidade e facilidade de uso, além de contribuir para a eficiência na especificação de sistemas IoT críticos.

Palavras–chave

Engenharia de Requisitos, Sistemas IoT Críticos, Safety e Security, Artefatos, STPA, Canvas

Abstract

Veiga, Ernesto. **SafeSecRETS: Integrated Support for Safety and Security Requirements Engineering in Critical IoT Systems**. Goiânia, 2025. 280p. PhD. Thesis. Programa de Pós-Graduação em Ciência da Computação, Instituto de Informática, Universidade Federal de Goiás.

[**Context**] The autonomy and connectivity characteristics inherent to critical IoT systems have required the joint treatment of safety and security from the early stages of design in order to prevent potential losses. [**Problem**] However, despite the growing attention in the literature to the integrated treatment of these requirements, there are still few research efforts that systematically address the activities and tasks of the Requirements Engineering (RE) process for safety and security, aiming to produce the documentation needed to guide subsequent phases of the development life cycle. [**Objective**] To address this gap, this research proposes artifacts to support the RE process for safety and security in critical IoT systems, covering project planning, elicitation, analysis, and specification of requirements. The proposal enables a co-analysis of safety and security requirements, allowing their interrelations to be addressed from the early stages of design. [**Methods**] The study followed the *Design Science Research* (DSR) methodology, which guided the design, construction, and evaluation of the artifacts, including: (i) a model for planning critical IoT projects (*SafeSecIoT Canvas*), instantiated from a metamodel for methodological support to the construction of canvas-based artifacts (*MM4Canvas*); (ii) a method that extends the System Theoretic Process Analysis (STPA) to enable the joint analysis of safety and security (*STPA-SafeSecIoT*); and (iii) a tool that integrates artifacts (i) and (ii) to support the activities and tasks of the RE process (*SafeSecRETS*). [**Results**] The evaluations conducted indicate that the proposed artifacts effectively support the execution of RE activities and tasks for safety and security, demonstrating high perceived usefulness and ease of use, and contributing to the efficiency of specifying critical IoT systems.

Keywords

Requirements Engineering, Critical IoT Systems, Safety and Security, Artifacts, STPA, Canvas

Contents

List of Figures

List of Tables

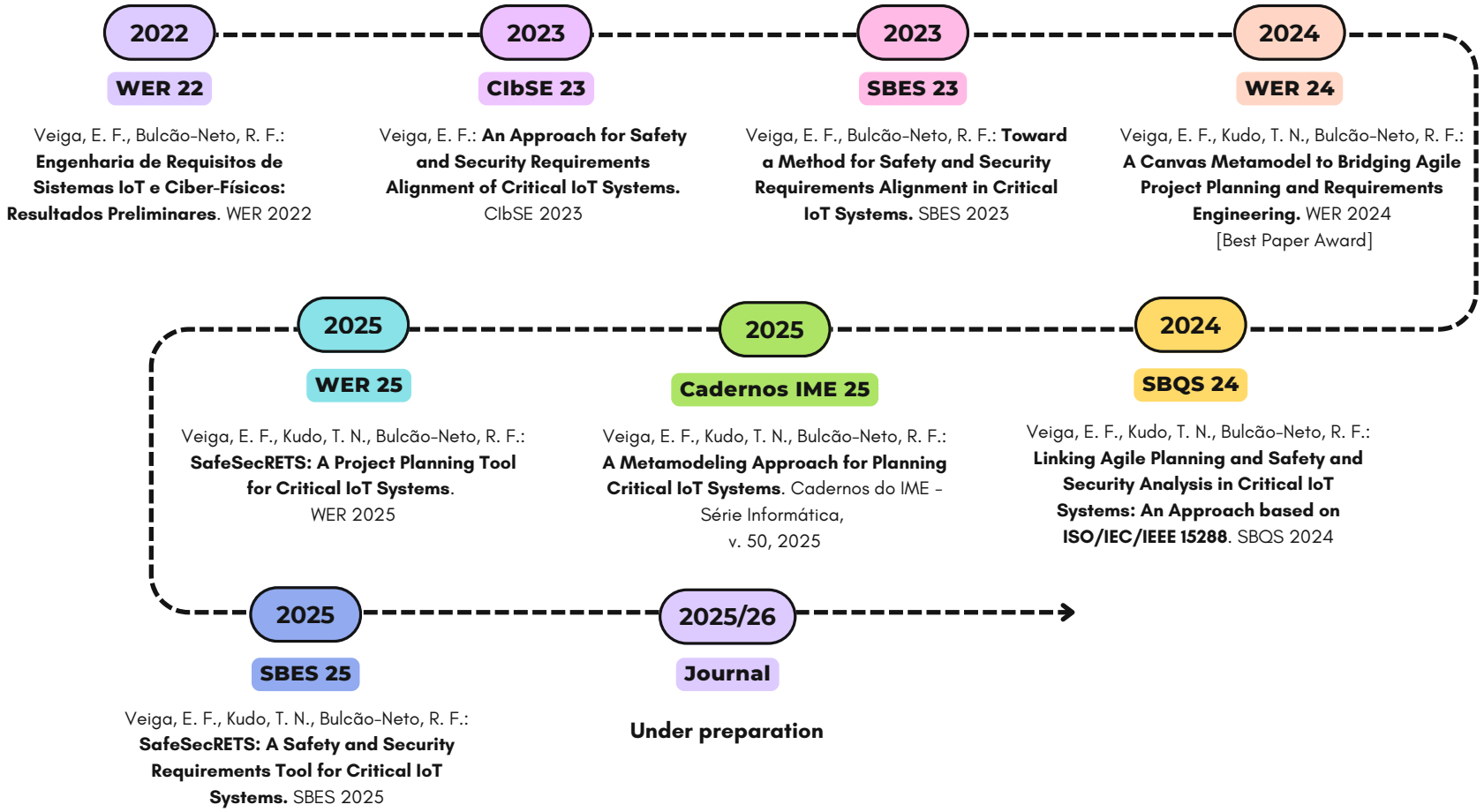
List of Acronyms

ACID	Atomicity, Consistency, Isolation, Durability
C	Component
CA	Control Action
CPS	Cyber-Physical Systems
DSR	Design Science Research
RE	Requirements Engineering
H	Hazard
IoT	Internet of Things
L	Loss
R	Responsibility
FHA	Functional Hazard Analysis
FMEA	Failure Mode and Effect Analysis
FR	Functional Requirements
FTA	Fault Tree Analysis
GQM	Goal Question Metric
NFR	Non-Functional Requirements
Req-Saf	Safety Requirement
Req-Sec	Security Requirement
RFID	Radio-Frequency Identification
S	Scenario
SC-Saf	Safety Constraint
SC-Sec	Security Constraint
SLM	Systematic Literature Mapping
SLR	Systematic Literature Review
STAMP	System-Theoretic Accident Model and Processes
STPA	System-Theoretic Process Analysis
SUS	System Usability Scale
T	Threat
TAM	Technology Acceptance Model
TIC	Information and Communication Technology
TLX	Task Load Index
UCA	Unsafe Control Action
UCA-Saf	Unsafe Control Action – Safety
UCA-Sec	Unsecured Control Action – Security
UMUX	Usability Metric for User Experience

List of Publications

List of publications resulting from this research to date:

- Veiga, E.F., Kudo, T.N., Bulcão-Neto, R.F.. 2025. **SafeSecRETS: A Safety and Security Requirements Tool for Critical IoT Systems**. In XXXIX Brazilian Symposium on Software Engineering (SBES '25), Brazil. https://doi.org/10.5753/sbes.2025.10843
- Veiga, E.F., Lima, K.L.S., Kudo, T.N., Bulcão-Neto, R.F.. 2025. **SafeSecRETS: A Project Planning Tool for Critical IoT Systems**. In 28th Workshop on Requirements Engineering (WER2025), Brazil. https://doi.org/10.29327/1588952.28-22
- Veiga, E.F., Kudo, T.N., Bulcão-Neto, R.F.. 2025. **A Metamodeling Approach for Planning Critical IoT Systems.** Cadernos do IME - Série Informática, 50 (jan. 2025). https://doi.org/10.12957/cadinf.2024.87931
- Veiga, E.F., Kudo, T.N., Bulcão-Neto, R.F.. 2024. **Linking Agile Planning and Safety and Security Analysis in Critical IoT Systems: An Approach based on ISO/IEC/IEEE 15288.** In Proceedings of the XXIII Brazilian Symposium on Software Quality (SBQS '24). Association for Computing Machinery, New York, NY, USA, 81–91. https://doi.org/10.1145/3701625.3701648
- Veiga, E.F., Kudo, T.N., Bulcão-Neto, R.F.. 2024. **A Canvas Metamodel to Bridging Agile Project Planning and Requirements Engineering.** In 27th Workshop on Requirements Engineering (WER2024), Argentina. https://doi.org/10.29327/1407529.27-18 (*Best Paper Award in the Regular Research Track*).
- Veiga, E.F., Bulcão-Neto, R.F.. 2023. **Toward a Method for Safety and Security Requirements Alignment in Critical IoT Systems.** In Proceedings of the XXXVII Brazilian Symposium on Software Engineering (SBES '23). Association for Computing Machinery, New York, NY, USA, 452–457. https://doi.org/10.1145/3613372.3613373
- Veiga, E.F.. 2023. **Uma Abordagem para Alinhamento de Requisitos de Segurança e Proteção de Sistemas IoT Críticos.** In Anais do XXVI Congresso Ibero-Americano em Engenharia de Software (CIbSE '23), Montevideo, Uruguai. SBC, Porto Alegre, Brasil, 277-284. DOI: https://doi.org/10.5753/cibse.2023.24712
- Veiga, E.F., Bulcão-Neto, R.F.. 2022. **Engenharia de Requisitos de Sistemas IoT e Ciber-Físicos: Resultados Preliminares.** In 25th Workshop on Requirements Engineering (WER2022), Brazil. https://doi.org/10.29327/1298262.25-11



Timeline of publications resulting from the research.

CHAPTER 1

Introduction

This chapter introduces the doctoral research documented in this thesis. It begins by providing the context of the study, highlighting the role and importance of safety and security requirements in the development of critical systems and their growing relevance in the Internet of Things (IoT) domain, which justifies and motivates this research. The research problem is then presented, grounded in the gaps identified in the literature, followed by the objectives, research questions, and an overview of the proposed solution. Finally, the chapter outlines the methodological procedures that guided the development of this work and describes the structure of the thesis.

1.1 Context and Rationale

Critical systems are those whose failures can lead to severe consequences, including harm to human life, environmental damage, or major economic and social disruption [Knight, 2002]. Such systems are increasingly common in domains such as healthcare, transportation, aerospace, and energy, where reliability is a non-negotiable requirement [Mailloux et al., 2019, Ajayi et al., 2025]. In these contexts, ensuring reliable operation under diverse and uncertain conditions is essential to preserving trust, safeguarding safety, and maintaining the continuity of operations [Ross et al., 2019].

Within these systems, two fundamental perspectives of dependability are particularly relevant: safety and security [Lisova et al., 2019a, Lyu et al., 2019]. Safety refers to the system's ability to operate without causing accidental harm to people or the environment, addressing risks associated with unintentional failures. Security, in contrast, concerns the protection of the system against intentional threats, focusing on risks arising from malicious attacks or unauthorized access.

Historically, safety and security have been treated separately for a long time, each with its own methods, standards, and communities [Line et al., 2006]. However, the growing interconnectivity and complexity of today's critical systems highlight their interdependence and the need for joint treatment from the beginning of the development life cycle [Wolf and Serpanos, 2018]. In these systems, an isolated approach to safety and se-

curity requirements is insufficient: a successful cyberattack can affect safety requirements and directly endanger people and other critical assets, while a safety-related malfunction can expose vulnerabilities that can be exploited to compromise the system's security requirements [Kriaa et al., 2015].

The advent and continued advancement of the IoT further amplifies these challenges. IoT systems are characterized by large-scale connectivity, device heterogeneity, continuous data exchange, and dynamic environments [Wolf and Serpanos, 2018, Silva et al., 2020]. These characteristics make IoT systems inherently complex, introducing multiple points of interaction, control, and potential failure [Xing, 2021]. The scale and diversity of IoT ecosystems also increase the difficulty of predicting behavior and controlling risks across the system [Rose et al., 2015, Nguyen-Duc et al., 2019].

When IoT technologies are applied to critical domains, the risks become even greater. As the attack surface expands, the possibility of cascading failures raises safety and security risks [Xing, 2021]. In this context, ensuring reliability requires methods that not only capture failures and threats but also consider the complex interactions and emergent properties that characterize critical IoT-based systems, mainly due to the autonomy that these systems exercise in relation to the environments, people, and other assets that participate in their operating context [Li et al., 2015, Yu et al., 2021].

Traditionally, safety analyses in critical systems have been based on well-established techniques such as *Failure Mode and Effect Analysis* (FMEA), *Fault Tree Analysis* (FTA), and *Functional Hazard Analysis* (FHA). These approaches have proven effective in systems with relatively stable architectures and well-defined failure modes [Yu et al., 2021]. However, they are inherently limited in their ability to model dynamic interactions, adaptive behaviors, and systemic interdependencies. As systems become more interconnected and operate in uncertain environments, the adequacy of these classical approaches is increasingly questioned [Leveson, 2016].

In response to these limitations, [Leveson and Thomas, 2018] proposed the *System-Theoretic Process Analysis* (STPA) method. This approach emerged as a modern risk analysis technique based on systems theory [Leveson, 2016]. A system can be understood as a union of interacting components, including software, hardware, people, processes, and data that operate in complex and interdependent ways [Leveson, 1995]. In STPA, safety is conceptualized as an emergent property that arises from the dynamic interactions between the components of the system. This systems theory-based approach offers a broader and more integrated perspective than its predecessors, making it particularly well-suited for analyzing the complexity of modern systems [Leveson, 2016, Leveson and Thomas, 2018].

In contrast to traditional methods, STPA adopts a top-down analysis perspective, focusing on control structures, feedback loops, and possible unsafe interactions between

system components and constraints that need to be enforced [Leveson, 2016]. This makes it more suitable for capturing the complex dynamics of contemporary systems, including those based on IoT, and even other emerging properties, such as security, necessary for these systems. By framing safety as a matter of inadequate control rather than isolated failures, STPA offers a more holistic and forward-looking approach to risk analysis.

Applying systems theory as a foundation, recent research has expanded the scope of STPA beyond safety analysis to also address the security needs of the system and seek a unified framework for analyzing hazards and threats [Friedberg et al., 2017, Veiga and Bulcão Neto, 2023]. These extensions aim, for example, to allow analysts to examine how inadequate control actions can create unsafe conditions and, at the same time, expose the system to potential attacks [Veiga and Bulcão Neto, 2023]. Thus, STPA extensions contribute to bridging the gap between safety and security, reinforcing their interdependence in the development of critical systems.

Although STPA and its extensions offer solid theoretical foundations for safety and security analysis in complex systems, their application often occurs in isolation, separate from a broader and more structured engineering process. In practice, this means that the results of STPA analyses are not always systematically integrated into an RE process, leading to gaps in traceability and loss of important information. As a consequence, the practical use of STPA-based approaches in supporting safety and security RE remains limited [Veiga et al., 2024a]. This highlights the need for systematic processes and support mechanisms that can operationalize the results of STPA application, enabling its use in critical IoT system development environments.

1.2 Problem and Motivation

At the beginning of this research, we conducted a Systematic Literature Mapping (SLM) to investigate and understand how the RE process has been conducted in the development of IoT systems [Veiga and Bulcão-Neto, 2022]. The results indicated that the most relevant types of requirements for this type of system were safety, security, efficiency, and reliability. According to [Chung and do Prado Leite, 2009], RNFs must be identified and addressed from the early stages of system design, since subsequent development activities lose meaning without a clear understanding of the real problems to be solved and the requirements to be met.

In addition to the SLM performed, reference works in this area indicate that approaches that integrate safety and security analysis have received increasing attention in recent years and highlight existing gaps and challenges:

- [Lisova et al., 2019a] present a Systematic Literature Review (SLR) on methods for co-analyzing safety and security, exploring trends and approaches for jointly

addressing these requirements in systems engineering. The authors emphasize that safety and security can negatively influence each other, and an efficient analysis of their interactions can reduce the effort required to achieve safe systems. They conclude that more efforts are needed to develop co-analysis approaches in different application domains and evaluate their effectiveness.

- [Lyu et al., 2019] analyze existing approaches to risk assessment and management from a safety perspective and their integration with security. The results show that research on these requirements in IoT systems is still in its early stages and requires significant improvements. The main gaps identified in existing methodologies include: i) difficulties in distinguishing between incidents caused by accidental failures and malicious attacks; ii) challenges in dealing with the complexity of systems, often in real time and requiring dynamic risk assessment, iii) lack of approaches to resolve conflicts between security and safety; and iv) absence of unified metrics for these requirements.
- [Aguilar-Calderón et al., 2022] conducted a systematic literature mapping (SLM) on RE for IoT, demonstrating that security requirements significantly increase project complexity and costs. The study also identified that the most frequently addressed non-functional requirements in IoT projects include security, privacy, maintainability, performance, scalability, interoperability, and safety, all of which demand particular attention in systems with autonomous components. Furthermore, the authors highlight that safety and security requirements are often insufficiently addressed due to their broad scope, complexity, and interdependencies, thereby underscoring research opportunities for advances in this area.

Despite the recognized importance of integrating safety and security into critical IoT systems, there is still a lack of systematic approaches and availability of supporting artifacts for the RE process. Many approaches tend to treat safety and security in isolation, with limited guidance on how to address their requirements and interdependencies throughout the early stages of system development. As a result, requirements are often captured in an ad hoc manner, leading to gaps, inconsistencies, and overlooked risks that compromise the reliability and resilience of the system.

In this context, based on the results of [Veiga and Bulcão-Neto, 2022] and other gaps identified in the literature, the research problem addressed in this thesis is:

> *The lack of a systematic process and artifacts to support safety and security RE activities for critical IoT systems.*

This research is motivated by the need to establish a structured process for safety and security RE that addresses the needs of critical IoT systems. The proposed approach

is based on STPA as a reference method for hazard and threat analysis and for specifying safety and security requirements. At the same time, it introduces a canvas model as an artifact to support project planning, scope definition, and elicitation of initial system requirements. By systematically integrating safety and security into the RE process, the approach aims to ensure greater consistency, traceability, and completeness in capturing requirements for critical IoT systems.

In this sense, this research identified as a promising direction the integration between strategic planning, which supports the definition of the project scope and the elicitation of requirements, with analysis based on STPA [Veiga et al., 2024a]. The top-down nature of both strategic planning and STPA enables integration and alignment between practices, supporting stakeholders and/or the technical team to collaboratively and iteratively the identification of objectives, key requirements, system components and assets, constraints, project risks, and safety and security issues, among other essential information, even in the early stages of system development, and use this information to support and improve STPA-based analysis.

To combine a structured approach to strategic planning with a method of safety and security analysis, we see the possibility of structuring a safety and security RE process for critical IoT systems based on supporting artifacts [Veiga, 2023]. The approach proposed in this research uses project planning methodologies and techniques to support the RE process, enabling and supporting the definition of requirements in a rigorous manner aligned with the perspectives of the system and stakeholders.

In this context, canvas-based approaches emerge as a valuable complement to STPA-based analysis. Widely used in agile and strategic management, canvas models provide visual and collaborative means to structure complex information, align stakeholders, and support decision-making [Osterwalder and Pigneur, 2010, Finocchio-Júnior, 2013]. The application of canvas models to the project planning of critical IoT systems allows stakeholders to define and clarify priorities and ensure that IoT, safety, and security concerns are explicitly addressed from the outset of the project [Veiga et al., 2024b]. Furthermore, when integrated with STPA, canvas-based planning has the potential to support other activities in the RE process for critical IoT systems.

Thus, one of the central elements of this research is the use of a canvas-based project planning approach, integrated with an extension of the STPA method for safety and security in support of the RE process for critical IoT systems. Canvas models allow stakeholders to collaboratively structure the problem space and essential system requirements, and STPA-based analysis complements this process by supporting detailed analysis and specification of safety and security. This not only strengthens the practical application of STPA, but also aligns RE activities with agile and strategic planning practices, improving communication and stakeholder engagement.

The proposal presented in this thesis is also based on established standards and practices, particularly ISO/IEC/IEE 15288 [ISO/IEC/IEEE, 2023], which defines processes for systems lifecycle management. By aligning the proposed process and artifacts with international standards, the approach reinforces its relevance and applicability in real-world contexts, where compliance and certification are often mandatory. Together, these elements highlight the motivation behind this research: to provide integrated methodological support based on supporting artifacts for the safety and security RE of critical IoT systems.

1.3 Objectives

1.3.1 General Objective

The overall objective of this work is to develop and evaluate artifacts to support the safety and security RE process for critical IoT systems, considering project planning, elicitation, analysis, and specification activities.

1.3.2 Specific Objectives

To achieve this purpose, the following specific objectives are defined:

- Define an approach and identify the types of artifacts needed for the joint treatment of safety and security requirements, from system concept to specification.
- Propose and build artifacts to support safety and security RE activities for critical IoT systems, as well as evaluate their applicability in practical scenarios.
- Instantiate and apply the proposed approach in different domains of critical IoT systems in order to verify its adaptability and scope.
- Evaluate the usefulness, ease of use, functional suitability, and other contributions of the proposal in supporting the RE of critical IoT systems that require the joint treatment of safety and security requirements.

1.4 Research Questions

Based on what has been presented, and supported by literature reviews, including the MSL on the safety and security RE process in critical IoT systems, this doctoral research is guided by the following research questions:

- **RQ01.** How can project planning based on canvas models and safety and security analysis based on STPA support the activities of the RE process in critical IoT systems?

- **RQ02.** How can elements of a canvas-based model be integrated into an extension of the STPA method to support the joint RE of safety and security of critical IoT systems?
- **RQ03.** To what extent does the support of specialized artifacts improve the perceived usefulness, ease of use, and intention to use of the proposed process for safety and security RE in critical IoT systems?

RQ1 aims to establish the relationship between canvas-based planning and an STPA-based method as practical mechanisms for directly supporting the RE process. Complementarily, RQ2 aims to investigate the integration between canvas and STPA, seeking methodological maturity for the proposal. Finally, RQ3 connects the technical/methodological part to empirical evaluation. This question aims to assess the acceptance and practical applicability of the proposal.

1.5 Methodological Procedures

1.5.1 Research Classification

[Easterbrook et al., 2008] present key questions that support the selection of a research method appropriate to the nature of the work, spanning from philosophical considerations to practical aspects related to its application and evaluation. Guided by these questions, we situated our study within the methodological category most aligned with its objectives and characteristics.

In the early stages of a research project, it is essential to formulate exploratory questions aimed at understanding the phenomena under investigation and identifying meaningful distinctions that clarify our conceptual framework [Easterbrook et al., 2008]. Accordingly, the research questions guiding this study can be classified within the categories of Description and Classification.

1.5.2 Research Method and Steps

This doctoral research follows the engineering and empirical research methods proposed by [Glass, 1994], which are widely adopted in Software Engineering. The engineering method focuses on the systematic observation of existing solutions, the identification of limitations, and the iterative proposal and refinement of improvements or new solutions. Complementarily, the empirical method emphasizes the evaluation of the proposed solutions through empirical evidence, such as experiments, case studies, or surveys [Glass, 1994, Wohlin, 2014].

To structure the development of this research, the phases of the engineering method are adopted, presented in Figure 1.1.

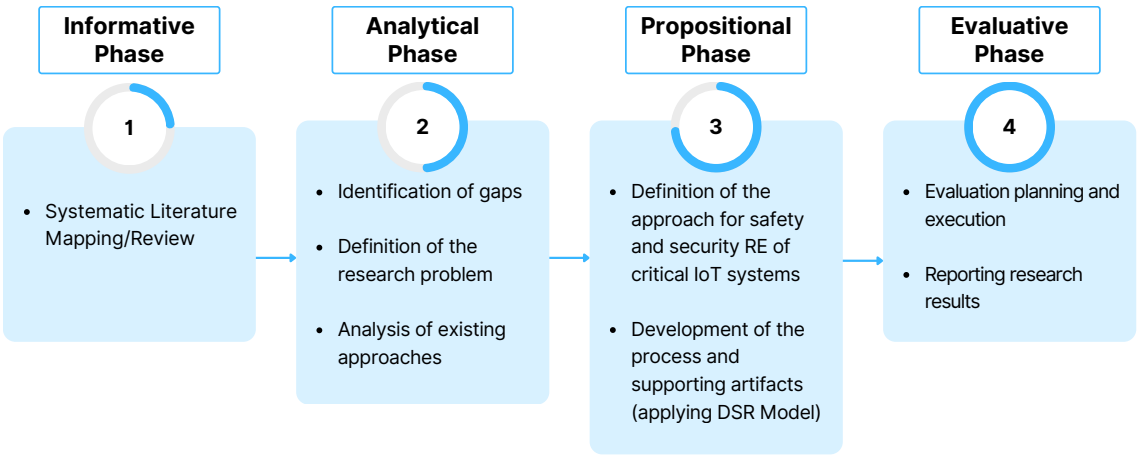


Figure 1.1: Phases of the research method and corresponding activities.

These phases are instantiated in the context of this thesis as follows:

- **Informational phase:** This phase focuses on the systematic acquisition and consolidation of knowledge. It comprises systematic literature reviews and mapping studies aimed at characterizing the state of the art in RE for safety and security in critical IoT systems, identifying existing approaches, artifacts, and tools, as well as their limitations.
- **Analytical phase:** This phase involves the critical analysis of the evidence gathered in the informational phase. It includes the identification of research gaps and open challenges, the analysis of existing safety and security RE approaches, and the definition of the research problem, objectives, and research questions, with particular attention to the joint treatment of safety and security in IoT systems.
- **Propositional phase:** This phase is dedicated to the design and development of a solution to address the identified problem. It encompasses the definition of a novel approach for safety and security requirements engineering in critical IoT systems, including the specification of a structured RE process, the development of supporting artifacts, and the implementation of a software tool, following the Design Science Research (DSR) paradigm.
- **Evaluative phase:** The evaluative phase aims to assess the proposed approach and artifacts in terms of relevance, utility, and quality. It includes the planning and execution of an evaluation with domain experts, focusing on perceived usefulness, ease of use, and adequacy for supporting safety and security RE in critical IoT systems. The results provide empirical evidence to support the discussion of contributions, limitations, and future research directions.

1.5.3 Applying Design Science Research

This research adopts the Design Science Research (DSR) methodology [Hevner et al., 2004, Peffers et al., 2007], structured according to the DSR Model [Pimentel et al., 2020]. This approach guides the systematic creation and evaluation of artifacts (such as models, methods, and tools) with the aim of solving practical problems in specific contexts. In this work, the focus is on supporting the RE process in critical IoT systems.

As shown in Figure 1.2, in the DSR Model, the design and construction of an *artifact* is guided by *behavioral conjectures* derived from a *theoretical framework* to address a *problem in context*, leveraging current knowledge: the state of the art and the state of the practice. Through *empirical evaluation*, the use of the artifact allows an assessment of the problem resolution and the validity of the underlying conjectures. Consequently, technical and scientific knowledge is produced through the design and scrutiny of the artifact's application, promoting advances in their respective domains.

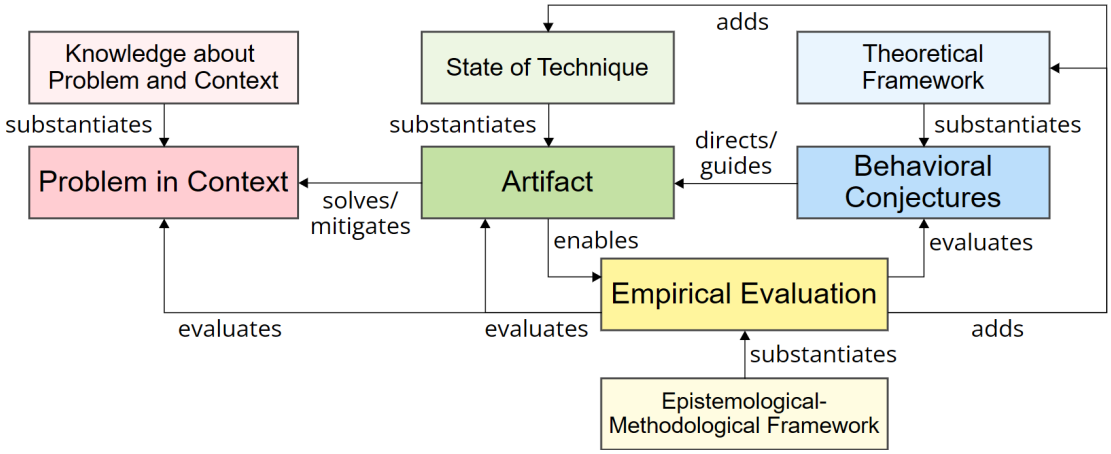


Figure 1.2: DSR Model [Pimentel et al., 2020].

Below, we describe the implementation of the DSR Model in our research.

Problem and Context

The lack of alignment between safety and security requirements can result in increased costs and even project failure. These challenges become even more significant as systems grow in complexity, as in the case of IoT [Sadvandi et al., 2012, Lisova et al., 2019b]. In addition, these requirements influence each other, sometimes adversely. When treated in disconnected processes, safety and security requirements can generate inconsistencies, conflicts, critical failures, and unacceptable losses [Kavallieratos et al., 2020a, Veiga, 2023]. In this context, it is necessary to investi-

gate integrated approaches that support the joint treatment of these requirements in order to reduce risks and increase system reliability.

Conjectures and Theoretical Framework

Based on the identified problem and the analyzed context, we formulate the following conjectures: i) the safety and security RE process can be performed more efficiently through specific techniques and artifacts that support its execution; ii) communication and collaboration between the different personas involved in the project are essential for the team to be able to integrate multiple perspectives (including IoT, safety, security, and the system's application domain); iii) breaking down the safety and security RE process into specific activities and tasks, supported by artifacts, helps reduce the complexity of the work to be done.

These conjectures are supported by the following theoretical framework: i) use of canvas integrated with STPA-based analysis, which provides mechanisms for defining the scope and analyzing the system, identifying hazards and threats, and deriving safety and security requirements in complex systems [Leveson and Thomas, 2018, Friedberg et al., 2017]; ii) the ISO/IEC/IEEE 15288:2023 standard, which defines system lifecycle processes and supports the integration of different perspectives and requirements [ISO/IEC/IEEE, 2023]; iii) agile methods, such as project planning with canvas, foster continuous communication and effective collaboration between technical team members and stakeholders [Finocchio-Júnior, 2013, Caroli, 2018, Veiga et al., 2024b]; iv) Model-Driven Engineering (MDE), which contributes to structuring and organizing the RE process, favoring abstraction and formalization of artifacts [Schmidt, 2006].

Artifacts and State of the Art

Based on the identified state of the art, the proposed artifacts were designed to address the problem in its specific context. Development took place in incremental cycles of the DSR Model, so that each cycle addressed particular aspects of the challenge of safety and security in critical IoT systems:

- First cycle: an extension of the STPA method, called *STPA-SafeSecIoT*, aimed at the integrated analysis and specification of safety and security requirements in critical IoT systems [Veiga and Bulcão Neto, 2023]
- Second cycle: a metamodel for canvas (*MM4Canvas*) to provide methodological support, and a specific model (*SafeSecIoT Canvas*) designed to support agile planning of critical IoT system projects by bridging technical and management perspectives [Veiga et al., 2024b, Veiga et al., 2024a];

- Third cycle: a software tool (*SafeSecRETS*), which integrates the artifacts proposed in the previous cycles, enabling their practical application in a unified and RE-oriented environment [Veiga et al., 2025b, Veiga et al., 2025a].

Throughout the development of the cycles presented, an approach was proposed and refined to link the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts, filling the gap between project planning and STPA-based analysis. Subsequently, this approach was also implemented in the *SafeSecRETS* tool.

1.6 Document Structure

This thesis is organized as follows:

- Chapter 2 presents the theoretical framework of the research, addressing the main topics involved in understanding and carrying out the proposal of this work;
- Chapter 3 presents the results of the MSL performed and the extension of this study with a focus on safety and security, in addition to discussing the main related works and the research gaps identified;
- Chapter 4 presents the *MM4Canvas* metamodel, proposed as methodological support for canvas development;
- Chapter 5 presents the *SafeSecIoT Canvas* model, an instance of *MM4Canvas* for project planning and elicitation of critical IoT system requirements, in addition to its integration with STPA-based security analysis approaches;
- Chapter 6 presents the *STPA-SafeSecIoT* method, an extension of STPA proposed for safety and security analysis and specification, also applied to critical IoT systems;
- Chapter 7 presents the evaluation of the *SafeSecIoT Canvas* artifact with undergraduate students from an IoT course. This evaluation served as an initial validation of the canvas-based project planning proposal;
- Chapter 8 presents the *SafeSecRETS* tool developed to implement the proposed artifacts and provide a collaborative software tool to support the RE process of critical IoT systems;
- Chapter 9 presents the evaluation of the *SafeSecIoT Canvas* model and the *STPA-SafeSecIoT* method (and their integration), as well as the *SafeSecRETS* tool, which implements both artifacts. The evaluation was conducted with experts who have extensive experience in RE, including safety and/or security;
- Chapter 10 presents final considerations, such as contributions, limitations, and lessons learned, in addition to discussing future work, concluding the report on the doctoral research conducted.

CHAPTER 2

Theoretical Foundations

This chapter introduces the key concepts that underpin the research. Section 2.1 defines Internet of Things (IoT) systems and their convergence with Cyber-Physical Systems (CPS). Section 2.2 outlines the fundamentals of safety and security and their role in requirements engineering for critical IoT systems. Section 2.3 provides an overview of the System-Theoretic Process Analysis (STPA) technique, which serves as a reference for extending safety analysis to security concerns. Finally, Section 2.4 presents project planning, the use of canvas and the main reference models adopted.

2.1 Internet of Things (IoT) Systems

Since the emergence of the term Internet of Things (IoT)[1] there has never been a consensus on its definition. Different views and definitions have been proposed and evolving since then, along with the technologies involved in its implementation and realization. Some time later, while the term IoT was still gaining momentum, a different community of experts came up with the definition of Cyber-Physical Systems (CPS), based on concepts of mechatronics, embedded systems, and pervasive computing. Despite their different origins, IoT and CPS show great convergence, as they refer to a related set of trends in the integration of digital resources with physical devices and engineering systems [Greer et al., 2019]. In addition, the evolution of these areas of research has brought their definitions closer together over time.

This section presents some consolidated definitions of IoT and CPS and also briefly discusses their convergence, which has led to the adoption of equivalence between the two terms. The purpose of these definitions and discussions is to describe the type of system to which the proposed alignment of requirements presented in this work applies.

[1]The term was first used by Kevin Ashton in a presentation on radio frequency identification at MIT in 1999 and in subsequent publications [Ashton et al., 2009]. For this reason, it is considered that the concept of IoT emerged from the RFID community and initially focused on the ability to track the location and status of any physical object or thing, mainly in supply chain applications.

2.1.1 Definitions and background

According to [Kopetz, 2011], the IoT is based on connecting physical things to the Internet, allowing access to data from remote sensors and remote control of the physical world. From this, the mixture of captured data with data retrieved from other sources, for example, data contained on the Web, gives rise to new synergistic services that go beyond the services that can be provided by an isolated embedded system. The idea of the IoT would be to interconnect the physical world with the digital world [Guth et al., 2016]. In this concept, sensors measure the parameters of the physical world, as well as changes in it, and this information is translated into data that can be processed by computers. In addition, based on this data, the IoT can act in the physical world through actuators.

The work of [Rajkumar et al., 2010], which considers CPS as the next computational revolution, argues that the role of devices has moved beyond merely connecting users to the Internet, evolving into a means of interconnecting the physical and cyber worlds and thus giving rise to CPS. In this context, [Borgia, 2014] defines CPS as the next generation of integrated ICT systems, in which computing and networking are embedded into physical processes to control and manage their dynamics, thereby enhancing efficiency, reliability, adaptability, and security.

According to [Minerva et al., 2015], a CPS is a system of collaborative computational elements that control physical entities. This is when mechanical and electrical systems (e.g., sensors and communication tools) embedded in products and materials are networked using software components. They use shared knowledge and process information to independently control logistics and production systems. Thus, CPS tends to go beyond mere unique identification and control of individual things to the level of networking between identified objects and sharing information about a specific condition to achieve a particular goal with better efficiency.

On the other hand, an IoT system starts at the level where a single "thing" is identified using a unique global identifier and can be accessed from anywhere, at any time. The level of information obtained when accessing the "thing" can be as low as static data that is stored in RFID tags. Primarily, IoT is concerned with the unique identification, Internet connection, and accessibility of "things." However, objects identified in an IoT system can still be networked to control a given scenario in a coordinated manner, in which case an IoT system can be considered to grow to the level of a CPS [Minerva et al., 2015].

In general, it can be said that a CPS is primarily concerned with the collaborative activity of sensors or actuators to achieve a specific goal, and to this end, the CPS uses an IoT system to perform the collaborative work of distributed systems [Minerva et al., 2015]. A CPS is a system of collaborative computational elements

that control physical entities. It is when mechanical and electrical systems are networked using software components. They use shared knowledge and process information to independently control logistics and production systems.

For [Miorandi et al., 2012], the term IoT is used as an umbrella keyword to cover various aspects related to the extension of the Internet and the Web into the physical realm, through the widespread deployment of spatially distributed devices with identification, detection, and/or actuation capabilities. Although originally aimed at identification and monitoring technologies, today the IoT also applies to the control of physical systems through systems integration [Gunes et al., 2014].

Thus, the terms IoT and CPS have distinct origins but overlapping definitions, both referring to trends in the integration of digital capabilities, including network connectivity and computing power, with physical devices and systems [Greer et al., 2019]. Examples of IoT and CPS systems range from smart vehicles to advanced industrial manufacturing systems, as well as applications in sectors as diverse as energy, agriculture, smart cities, among many others.

2.1.2 Convergence between the IoT and CPS

As mentioned earlier, the concepts of Cyber-Physical Systems (CPS) and the Internet of Things (IoT) emerged from different communities, with CPS arising mainly from an engineering and systems control perspective and IoT originating with the use of RFID [Greer et al., 2019]. However, in most academic and project activities, and in many of the proposed definitions, the difference between "Internet of Things" and "Cyber-Physical Systems" is unclear, and it is difficult to distinguish between the two terms [Minerva et al., 2015]. Many studies and authors consider the two definitions to be different explanations for the same idea and use the two terms interchangeably.

The National Institute of Standards and Technology (NIST)[2] conducts an in-depth analysis of IoT and CPS publication trends [Greer et al., 2019]. This study highlights three temporal phases in IoT publications: low numbers and growth rates between 2005 and 2009, an increase from 2010 to 2013, and rapid growth after 2014. The study's analysis of 30 IoT definitions reveals terminology from the network/information technology communities. Furthermore, analysis of these definitions over time reveals an evolution from trackable objects and data to hybrid systems in which these objects are components of interactive and intelligent systems, making recent IoT definitions largely interchangeable with those of CPS.

[2] A US federal institute that acts as a technology and metrology standards agency for the US government. Established in 1901, its goal is to promote innovation and industrial competitiveness by developing and promoting technical standards and guidelines for various areas.

Meanwhile, analysis of CPS publication trends over the past decade shows a pattern of steady expansion. The study by [Greer et al., 2019] reviewed 31 published CPS definitions and reveals common terminology from computer science and systems engineering. The definitions are fairly consistent over time and highlight a set of six common CPS characteristics: hybrid physical and logical systems, hybrid analytical and measurement methods, control, component classes, time, and reliability.

[Greer et al., 2019] also analyzed 11 publications comparing and contrasting IoT and CPS, and found that the distinctions between them are related to issues of control, platform, internet, and human interactions. However, according to the authors themselves, a more in-depth analysis indicates that these issues are insufficient to draw a reliable distinction between IoT and CPS. The lack of consistent distinctive metrics and the convergence of definitions indicate an emerging consensus around the equivalence of the concepts. This convergence creates opportunities for progress through the integration of research, innovation, and standardization efforts of the respective communities.

According to [Greer et al., 2019], the implications of a unified perspective on IoT and CPS include the opportunity for research communities to work together on developing new, unified discrete and continuous methods for the design, operation, and assurance of IoT and CPS; and highlight the importance of close cyber-physical linkage (e.g., robust sensing and actuation, secure systems, digital models, etc.) as the basis for the transformational nature of IoT and CPS concepts. The depth of these implications is illustrated in examples of design assurance and cyber-physical security for complex IoT/CPS systems. The latter example is directly related to this proposal.

2.1.3 Critical IoT Systems

Reinforcing the definition presented in Chapter 1, in the context of this work, we consider crtitical IoT systems those whose safety and/or security requirements are essential to avoid unacceptable losses. That is, in critical IoT systems, problems related to these requirements can lead to injury or death, environmental damage, unauthorized disclosure of information, loss of highly relevant data, severe financial damage, among other unacceptable outcomes, based on the definition of losses proposed by [Leveson and Thomas, 2018] and the concept of critical systems [Sommerville, 2016].

Critical IoT systems are characterized by their high degree of connectivity, heterogeneity of devices and technologies, and frequent integration with cyber-physical components. These systems often operate in dynamic and unpredictable environments, which makes them more vulnerable to both accidental failures and intentional attacks. They also handle sensitive information flows and must ensure properties such as availability, integrity, confidentiality, and resilience in the face of failures or intrusions.

Moreover, critical IoT systems are usually deployed in domains with strict regulatory or operational constraints, such as healthcare, transportation, energy, and industrial automation. In these contexts, their operation is closely tied to human safety, mission success, and societal trust. Consequently, their development requires not only advanced technical solutions but also systematic processes for RE, risk management, and assurance of safety and security properties.

2.2 *Safety* e *Security*

As computer systems become deeply embedded in businesses and personal lives, problems resulting from system and software failures are increasing [Sommerville, 2016]. For instance, a server software failure at an e-commerce company can cause significant loss of revenue and customers. A software error in a control or multimedia system embedded in a car may lead to costly recalls, causing severe damage to the automaker, and in the worst case, contributing to accidents and losses, including loss of life. Similarly, a malware infection in a company's computers demands costly cleanup operations and may result in the loss or compromise of confidential information.

Software-intensive systems have become increasingly important to governments, businesses, and individuals. As a result, there is a growing dependence on these systems. They must be available when required, and the system as a whole must function correctly, without undesirable side effects such as unauthorized disclosure of information, unexpected behavior, or improper actions. In summary, modern society requires software systems that are dependable.

2.2.1 Safety Fundamentals

Safety-critical systems are those in which it is essential that the system always operates safely [Sommerville, 2016]. In other words, it must never cause harm to people or the environment, regardless of whether or not the system complies with its specifications. Examples of safety-critical systems include control and monitoring systems in aircraft, process control systems in industries, and automotive control systems, among others.

The development of safety-critical systems works with hazard-oriented techniques, which consider possible system accidents that may occur [Sommerville, 2016]:

- Hazard prevention: The system is designed so that hazards are avoided. For example, a paper cutting system that requires the operator to use both hands to press separate buttons simultaneously avoids the risk of the operator's hands getting in the way of the blade.

- Detection and removal of hazards: The system is designed so that hazards are detected and removed before they result in an accident. For example, a chemical plant system can detect excessive pressure and open a relief valve to reduce the pressure before an explosion occurs.
- Damage limitation: The system may include protective features that minimize the damage that can result from an accident. For example, an aircraft engine typically includes automatic fire extinguishers. If a fire starts in the engine, it can often be controlled before it poses a threat to the aircraft.

A hazard is a state of the system that can lead to an accident. Hazards are not (or do not necessarily become) accidents: often dangerous situations can occur that are resolved or end without any problems. However, accidents are always preceded by hazards. Therefore, reducing hazards reduces accidents. The following is a vocabulary of key specialized terms used in safety-critical systems [Sommerville, 2016]:

- *Accident*: an unplanned event or sequence of events that results in losses such as human death or injury, damage to property, or damage to the environment.
- *Damage*: a measure of the loss resulting from an accident. Damage can range from many people being killed as a result of an accident to minor injuries or property damage.
- *Hazard*: a condition with the potential to cause or contribute to an accident.
- *Risk*: a measure of the likelihood that the system will cause an accident. Risk is assessed by considering the probability of a hazard, the severity of the hazard, and the probability that the hazard will lead to an accident.

As systems become increasingly complex, it becomes more difficult to understand the relationships between different parts of the system. Consequently, it is more difficult to predict the consequences of a combination of unexpected events or system failures [Sommerville, 2016]. For this reason, new techniques have been proposed for safety analysis in complex systems, for example, STPA [Leveson and Thomas, 2018].

2.2.2 Security Fundamentals

The widespread adoption of the Internet since the 1990s has introduced a persistent and evolving challenge for software engineers: the design and implementation of secure systems. As the number of Internet-connected systems has grown, so too has the diversity and sophistication of external attacks targeting them. Consequently, the complexity of producing secure systems has increased significantly. Systems engineers must account not only for threats posed by malicious and technically skilled attackers

but also for vulnerabilities introduced through accidental errors during the development process [Sommerville, 2016].

In the current context, it is essential to design systems capable of resisting external attacks and recovering from them. Without appropriate security precautions, networked systems are inevitably compromised. Attackers may exploit hardware resources, steal confidential information, or disrupt the services provided by the system. Within secure systems engineering, three key dimensions of security are particularly critical [Goodrich and Tamassia, 2011][Sommerville, 2016]:

1. *Confidentiality*: Information in a system may be disclosed or made accessible to unauthorized persons or programs. For example, the theft of credit card data from an e-commerce system is a confidentiality problem.
2. *Integrity*: Information in a system may be damaged or corrupted, rendering it incorrect or unreliable. For example, a worm that alters data in a system is an integrity problem.
3. *Availability*: Access to a system or its data, which is normally available, may be prevented. A denial-of-service attack that overloads a server is an example of a situation in which system availability is compromised.

In this sense, the term cybersecurity has been increasingly used in discussions about system security [Sommerville, 2016]. Cybersecurity is a very broad term that covers all aspects of security for citizens, businesses, and critical infrastructure against threats arising from the use of computers and the Internet. Its scope includes all levels of the system, from hardware and networks, through application systems, to mobile devices that can be used to access these systems.

Security is a system attribute that reflects the system's ability to protect itself against malicious attacks, whether internal or external. These external attacks are possible because most computers and mobile devices are networked and can therefore be accessed by outsiders. Examples of attacks include the installation of viruses, unauthorized use of system services, or unauthorized modification of a system or its data.

The basic specialized terminology associated with security comprises the following terms [Pfleeger et al., 2015]:

- *Asset*: something of value that must be protected. The asset may be the software system itself or the data used by that system.
- *Attack*: the exploitation of a vulnerability in a system, in which an attacker aims to cause damage to one or more assets of the system. Attacks can occur from outside the system (external attacks) or from authorized internal personnel (internal attacks).

- *Control*: a security measure that reduces the vulnerability of a system. Encryption is an example of a control that reduces the vulnerability of a weak access control system.
- *Exposure*: possible loss or damage to a computer system. This could be loss or damage to data, or it could be a loss of time and effort if recovery is necessary after a security breach.
- *Threat*: circumstances that have the potential to cause loss or damage to a system. You can think of a threat as a vulnerability in the system that is subject to attack.
- *Vulnerability*: a weakness in a computer-based system that can be exploited to cause loss or damage.

System vulnerabilities may arise from deficiencies in requirements, design, or implementation, as well as from human, social, or organizational failures [Sommerville, 2016]. Examples include the use of weak passwords, insecure storage of credentials in accessible locations, or errors by system administrators when configuring access control mechanisms or system files. Nevertheless, such issues cannot be attributed solely to human error. In many cases, user mistakes or omissions stem from inadequate system design choices. For example, policies that mandate frequent password changes (which encourage users to write down their credentials) or overly complex configuration mechanisms that increase the likelihood of misconfiguration.

Security must be a fundamental principle in the creation of reliable, available, and secure intensive software systems. It is not an add-on that can be added later, but must be considered at all stages of the development life cycle, from initial requirements to system operation [Sommerville, 2016].

Four types of threats to security may arise [Sommerville, 2016]:

1. *Interception threats* that allow an attacker to gain access to an asset. E.g., a situation where an attacker gains access to an individual patient's records.
2. *Interruption threats* that allow an attacker to render part of the system unavailable. E.g.: a denial-of-service attack on a system database server.
3. *Modification threats* that allow an attacker to tamper with a system asset. E.g.: an attacker alters or destroys a user's record.
4. *Fabrication threats* that allow an attacker to insert false information into a system. E.g.: in a banking system, false transactions can be added to the system that transfer money to the perpetrator's bank account.

Security must be a fundamental principle in the creation of reliable, available, and secure software-intensive systems. It is not an add-on that can be added later, but must be considered at all stages of the development life cycle, from initial requirements to system operation [Sommerville, 2016].

2.2.3 Safety and Security Requirements Engineering

Risk-based safety requirements specification is an approach used in critical systems engineering, where the risks faced by the system are identified and the requirements to avoid or mitigate those risks are identified. For safety-critical systems, this translates into a process driven by hazard identification. It can be used for all types of reliability requirements.

There are four activities in a hazard-oriented safety specification process, shown in Figure 2.1 [Sommerville, 2016]:

- *Hazard identification*: responsible for identifying hazards that may threaten the system. These hazards can be recorded in a formal document that records safety analyses and assessments and can be submitted to a regulator as part of a safety case.
- *Hazard assessment*: decides which hazards are the most severe and/or most likely to occur. These should be prioritized when deriving safety requirements.
- *Hazard analysis*: root cause analysis process that identifies events that could lead to a hazard occurring.
- *Risk reduction*: based on the results of the hazard analysis, leading to the identification of safety requirements. These requirements may be concerned with ensuring that a hazard does not arise or lead to an accident or that, if an accident occurs, the associated damage is minimized.

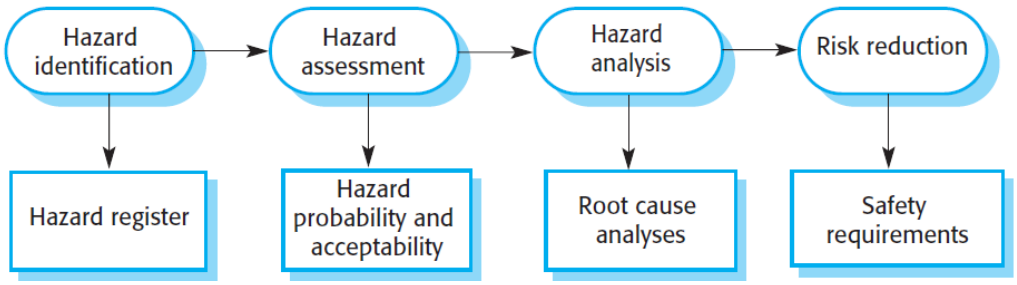


Figure 2.1: Process for identifying safety requirements [Sommerville, 2016].

The specification of system security requirements has much in common with the specification of safety requirements. It is not possible to specify safety or security requirements as probabilities: like safety requirements, security requirements are generally "must not" requirements that define unacceptable system behavior rather than required functionality. However, security can be an even more challenging problem than safety, for several reasons [Sommerville, 2016]:

- When considering safety, it can be assumed that the environment in which the system is installed is not hostile, i.e., no one is attempting to cause a safety-related

incident. However, when considering security, it is necessary to assume that attacks on the system are deliberate and that the attacker may be aware of the system's weaknesses.

- When system failures occur that pose a risk to safety, the errors or omissions that caused the failure are sought. When deliberate attacks cause system failure, finding the root cause can be more difficult, as the attacker may try to hide the cause of the failure.
- It is generally acceptable to shut down a system or degrade system services to prevent a safety-related failure. However, attacks on a system can be denial-of-service attacks, whose goal is to compromise the availability of the system. Shutting down the system means that the attack was successful.
- Safety-related events are accidental and not created by an intelligent adversary. On the other hand, an attacker may probe a system's defenses in a series of attacks, modifying the attacks as they learn more about the system and its responses.

These distinctions between safety and security mean that security requirements are generally broader than safety requirements, as they also involve external factors. Many types of security requirements cover the different threats faced by a system. In this sense, [Firesmith et al., 2003] identified 10 types of security requirements that can be included in a system specification:

1. *Identification requirements* specify whether or not a system must identify its users before interacting with them.
2. *Authentication requirements* specify how users are identified.
3. *Authorization requirements* specify the privileges and access permissions of identified users.
4. *Immunity requirements* specify how a system should protect itself against viruses, *worms*, and similar threats.
5. *Integrity requirements* specify how data corruption can be prevented.
6. *Intrusion detection requirements* specify which mechanisms should be used to detect attacks on the system.
7. *Non-repudiation requirements* specify that a party to a transaction cannot deny its involvement in that transaction.
8. *Privacy requirements* specify how data privacy should be maintained.
9. *Security audit requirements* specify how system usage can be audited and verified.
10. *System security maintenance requirements* specify how an application can prevent authorized changes from accidentally overriding its security mechanisms.

Obviously, not all of these types of security requirements appear in all systems. The particular requirements depend on the type of system, the usage situation, and

the expected users. Preliminary risk assessment and analysis aim to identify generic security risks to a system and its associated data. This risk assessment is an important input to the security requirements engineering process. Security requirements can be proposed to support general risk management strategies for prevention, detection, and mitigation [Sommerville, 2016]:

- *Risk prevention requirements* establish the risks that must be avoided when designing the system so that these risks simply cannot arise.
- *Risk detection requirements* define mechanisms that identify the risk, should it arise, and neutralize the risk before losses occur.
- *Risk mitigation requirements* establish how the system should be designed so that it can recover and restore system assets after a loss has occurred.

The following describes a generic process for identifying risk-based security requirements, presented in Figure 2.2 [Sommerville, 2016]. The steps in the process are:

1. *Identification of assets*: the system assets that may require protection are identified. The system itself or specific system functions can be identified as assets, as well as data associated with the system;
2. *Asset value assessment*: where the value of each identified asset is estimated;
3. *Exposure assessment*: assessment of potential losses associated with each asset. This process should take into account direct losses, such as information theft, recovery costs, and possible loss of reputation;
4. Threat identification: identification of threats to system assets;
5. *Attack assessment*: each threat is broken down into attacks that can be made on the system and the possible ways in which these attacks can occur;
6. *Control identification*: controls that can be implemented to protect an asset are proposed. Controls are technical mechanisms, such as encryption, that you can use to protect assets;
7. *Feasibility assessment*: the technical feasibility and costs of the proposed controls are assessed. It is not worth having expensive controls to protect low-value assets;
8. *Definition of security requirements*: knowledge about exposure, threats, and control assessments is used to derive the security requirements of the system. These requirements may apply to the system infrastructure or to an application.

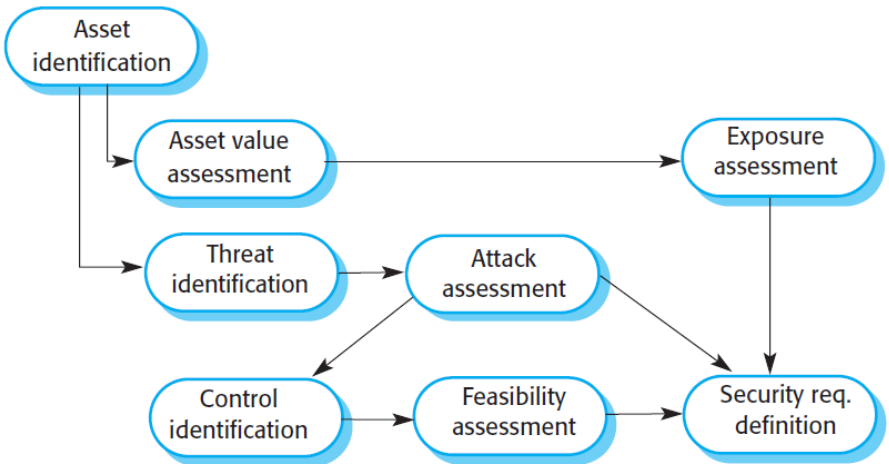


Figure 2.2: Process for identifying security requirements [Sommerville, 2016].

2.3 System-Theoretic Process Analysis (STPA)

System-Theoretic Process Analysis (STPA) is a hazard analysis technique/method based on an extended accident cause model called STAMP [Leveson, 2016]. In addition to component failures, STPA assumes that accidents can also be caused by unsafe interactions of system components, where none of them necessarily need to have failed. Some of the advantages of STPA over traditional hazard/risk analysis techniques are [Leveson and Thomas, 2018]:

- *Highly complex systems can be analyzed*. "Unknown unknowns" that were previously found only in operations can be identified early in the development process and eliminated or mitigated. Intended and unintended functionalities are addressed.
- *Unlike traditional risk analysis methods, STPA can be initiated in the initial concept analysis to assist in identifying safety requirements and constraints*. These can then be used to design safety (and security) into the system architecture and design, eliminating the costly rework involved when design flaws are identified late in development or during operations. As the design is refined and more detailed design decisions are made, the STPA analysis is also refined to help make increasingly detailed design decisions. Complete traceability of requirements to all system artifacts can be easily maintained, improving maintainability and system evolution.
- *STPA includes software and human operators in the analysis*, ensuring that the risk analysis includes all potential causal factors in losses.
- *STPA provides documentation of system functionality*, which is often missing or difficult to find in large, complex systems.
- *STPA can be easily integrated with other system engineering processes and model-based systems engineering*.

Since its creation, many evaluations and comparisons of STPA have been made in relation to other more traditional risk analysis methods, such as fault tree analysis (FTA), failure modes and effects criticality analysis (FMECA), event tree analysis (ETA), and hazard and operability analysis (HAZOP). In all of these evaluations, STPA identified all of the causal scenarios found by traditional analyses, and also identified several others, often related to software and without failures: scenarios that traditional methods were unable to find. In some cases, where there was an accident that was not reported to analysts, only STPA found the cause of the accident. In addition, STPA proved to be much less time-consuming and resource-intensive than traditional methods [Leveson and Thomas, 2018].

2.3.1 STPA Overview

STPA is divided into four steps, described below and presented in Figure 2.3 [Leveson and Thomas, 2018]:

- **Define the purpose of the analysis**. This is the first step in any analysis method. It seeks to answer the following questions, among others:
  - What types of losses does the analysis aim to prevent?
  - What is the system to be analyzed and what is the system boundary?
- **Model the control structure**. This consists of building a model of the system called a control structure. A control structure captures functional relationships and interactions by modeling the system as a set of feedback control loops. The control structure usually starts at a more abstract level and is refined iteratively to capture more details about the system.
- **Identify unsafe control actions**. This consists of analyzing the control actions in the control structure to examine how they can lead to the losses defined in the first step. These unsafe control actions are used to create functional requirements and constraints for the analyzed system.
- **Identify loss scenarios**. The fourth step identifies the reasons why unsafe control can occur in the system. The scenarios are created to explain: i) how incorrect feedback, inadequate requirements, design errors, component failures, and other factors can cause unsafe control actions and ultimately lead to losses; and ii) how safe control actions may be provided but not followed or executed properly, leading to a loss.

Once scenarios are identified, they can serve multiple purposes: supporting the definition of additional requirements, guiding the selection of countermeasures, informing system architecture, and driving design recommendations or new design decisions (when STPA is applied during development). When applied after project completion, scenarios

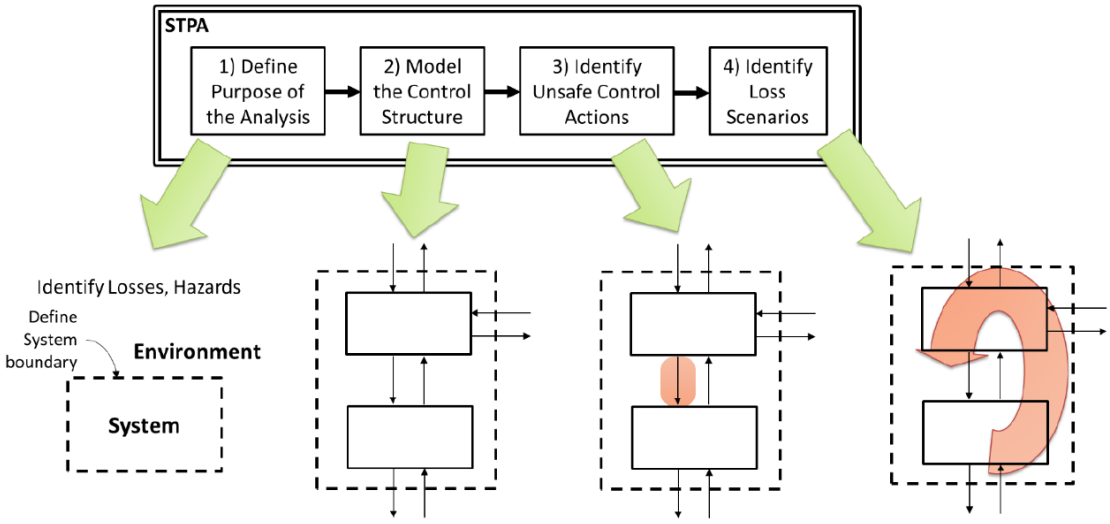


Figure 2.3: Overview of the STPA method [Leveson and Thomas, 2018].

can be used to evaluate or revisit existing design decisions and identify gaps. In both cases, they also support the definition of test cases and test plans, the development of risk indicators, and other analysis activities.

2.3.2 Steps in safety analysis with STPA

This subsection details the steps of STPA, based on the *STPA Handbook* [Leveson and Thomas, 2018].

Step 1: Define the Purpose of the Analysis

The first step in applying STPA is to define the purpose of the analysis, which can be divided into four parts: 1) identify losses, 2) identify hazards at the system level, 3) identify constraints at the system level, and 4) refine hazards (optional). An overview of this step is presented in Figure 2.4.

The main definitions for understanding the activities carried out in this stage are [Leveson and Thomas, 2018]:

- *Loss*: A loss involves something of value to stakeholders. Losses may include loss of life or human injury, property damage, environmental pollution, mission loss, loss of reputation, loss or leakage of sensitive information, or any other loss that is unacceptable to stakeholders.
- *Hazard*: a state of the system or set of conditions that, together with a particular set of worst-case environmental conditions, will lead to a loss.
- *System*: a set of components that act together as a whole to achieve some common goal, objective, or end. A system may contain subsystems and may also be part of a larger system.

- *System-level constraint*: A system-level constraint specifies the conditions or behaviors of the system that must be met to avoid hazards (and ultimately avoid losses).

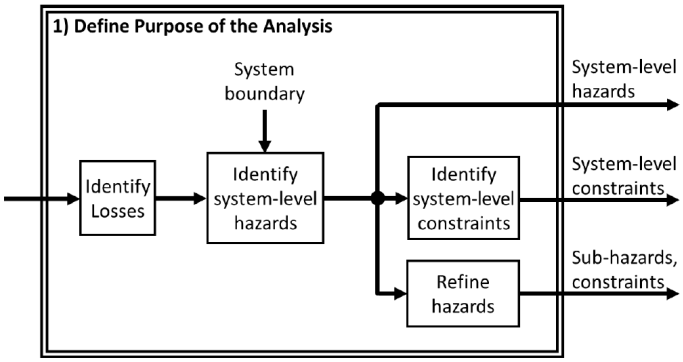


Figure 2.4: Overview of the first stage of STPA: defining the purpose of the analysis [Leveson and Thomas, 2018].

Once the hazards at the system level have been identified, it is easy to identify the constraints at the system level that must be imposed: simply reverse the condition. Thus, we have that:

<Hazard> = <System> & <Unsafe Condition> & <Link to Losses>
<System Level Constraint> = <System> & <Condition to Apply> & <Link to Hazard>

System-level constraints can also define how the system should minimize losses in the event of hazards occurring. This type of constraint can be written using the following syntax:

<System Level Constraint> = If <hazard> occurs, then <describe what needs to be done to prevent or minimize a loss> & <Link to Hazard>

The remainder of the STPA analysis will systematically identify scenarios that may violate these constraints, leading to system-level risks and losses.

Step 2: Model the Control Structure

A hierarchical control structure is a system model composed of feedback control loops. An effective control structure will impose constraints on the behavior of the system as a whole. The following are important concepts for understanding a control structure and details of how this approach is implemented.

Important concepts for understanding a control structure

Control is a fundamental concept in systems engineering, particularly in the context of complex systems [Leveson and Thomas, 2018]. It does not imply enforcing strict compliance with predefined functions or procedures; rather, it refers to mechanisms that

ensure system objectives and constraints are met. Control can be achieved through design features (e.g., interlocks and fail-safe mechanisms), process controls (e.g., maintenance and manufacturing procedures), or other strategies. Without some form of control, it would be impossible to guarantee that system goals are satisfied.

A control structure models the system in terms of control loops and feedback mechanisms. An effective control structure constrains system behavior to maintain safety and performance. Typically, these structures are hierarchical, comprising multiple interrelated control loops [Leveson and Thomas, 2018]. In systems engineering, the control loop is a primary means of implementing control [Leveson, 2016]. A control system directs, regulates, or commands the behavior of processes or devices by means of such loops. Applications range from simple domestic systems (such as thermostats regulating home temperature) to large-scale industrial control systems that govern complex processes and machinery.

Figure 2.4 shows the basic structure of a generic control loop.

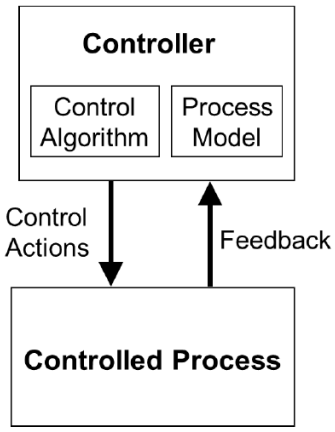


Figure 2.5: Generic control loop [Leveson and Thomas, 2018].

In general, a controller can provide control actions to control some process and impose restrictions on the behavior of the controlled process. The control algorithm represents the controller's decision-making process, determining the control actions to be provided. Controllers also have process models that represent the internal rules of the controller used to make decisions. Process models may include rules about the process being controlled or other relevant aspects of the system or environment. Process models can be updated in part by feedback used to observe the controlled process.

The components of a control loop are:

- The controller: to control a process or system, a controller is required; this is the component that can affect the behavior of the controlled process or system, in this case, keeping the system within safety and security limits by satisfying restrictions on system operation, which are controlled through control actions.

  - The process model: the controller must contain some model of the state of the controlled process, that is, a process model, which is a view of the state of the system: a set of properties with their determined values at a given moment.
  - The control algorithm: an element that represents the controller's decision-making process, which determines the control actions that will be performed on the system being controlled.
- Control actions: these are actions used by controllers according to the state of the system to control a process.
- Feedback: an element that provides information to the controller about the current state of the controlled process, i.e., about the state of the system at a given moment.
- The controlled process: is the process/system being controlled by the controller, whose states will be affected by the control actions issued by a controller; other elements may affect a controlled process, such as: other controllers, process inputs from the environment, other inputs or disturbances, among others.

A control structure is a functional model rather than a physical one, in which connections represent commands (control actions) and feedback, but do not necessarily correspond to physical links. Moreover, it is not an executable model, particularly because it may include human agents. Nevertheless, the STPA technique can derive requirements and other outputs from the control structure that can later support the development of executable models and specifications.

Importantly, a control structure does not assume compliance: the transmission of a control action does not guarantee that it will be executed, just as the provision of feedback cannot always be assumed. Instead, control actions and feedback represent the intention to establish mechanisms during system design for exchanging this information, without prescribing how controllers will actually behave in practice.

Obviously, most systems typically have multiple overlapping and interactive control loops. Multiple interactive control loops can be modeled in a hierarchical control structure, as shown in Figure 2.6.

Modeling the control structure

Control structure modeling begins with an abstract control structure and adds details iteratively. In many cases, the control structure and control loops within the system may be obvious or may be reused from previous applications. For more complex control structures, one way to start is to identify the basic subsystems needed to enforce the constraints and prevent the hazards identified earlier. The control structure can be refined by defining how each subsystem will be controlled.

During the development of the control structure, responsibilities can be assigned to each entity in the control structure. These responsibilities are a refinement of the

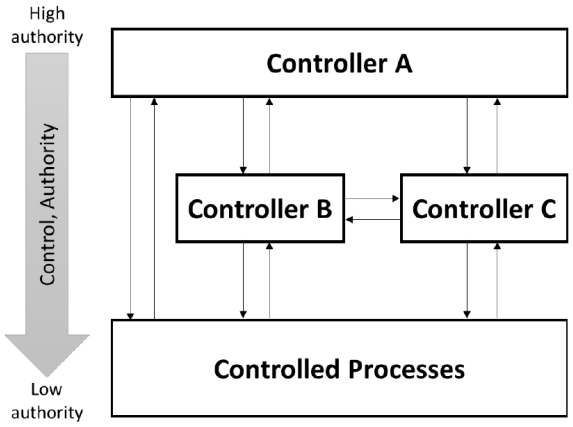


Figure 2.6: Overlapping and interactive control loops with different levels of authority [Leveson and Thomas, 2018].

system-level constraints, i.e., what each entity needs to do so that together the system-level constraints are enforced.

Step 3: Identify Unsafe Control Actions

Once the control structure has been modeled, the next step is to identify unsafe control actions, as shown in Figure 2.7 [Leveson and Thomas, 2018].

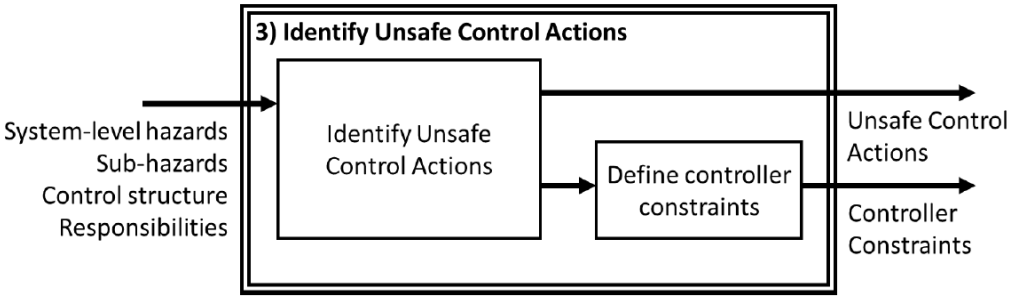


Figure 2.7: Overview of the third stage of STPA: analysis of unsafe control actions [Leveson and Thomas, 2018].

An Unsafe Control Action (UCA) is a control action that, in a particular context and worst-case environment, will lead to a hazard. Each UCA must specify under what conditions (in what context) the control action is unsafe. By identifying such unsafe actions, it is possible to subsequently eliminate these instances from the system design or find ways to mitigate them. Any relevant context can be referenced in a UCA, including environmental conditions, states of controlled processes, states of the controller, previous actions of the controller (e.g., repetitive actions), states of other controllers, previous actions of others, simultaneous or conflicting actions, parameters or properties of the control action, or any other relevant conditions. Using words such as “when,” “while,” or “during” in the construction of the UCA is often helpful in developing the context.

A UCA contains five parts:

<Unsafe Control Action> = <Source> <Type> <Control Action> <Context> <Link to Hazards>

The first part of a UCA is the *controller* that can provide the control action. The second part is the *type of unsafe control action* (provided, not provided, too early or too late, interrupted too early, or applied for too long). The third part is the *control action or command* itself (from the control structure). The fourth part is the *context* discussed above, and the last part is the *link to hazards* (or sub-hazards). UCAs are usually written with each part in the same order shown above, but in some cases it may be clearer or more natural to use a different order. The order is not critical. The key point is that UCAs must contain these five parts.

Once the UCAs have been identified, they can be translated into constraints on the behavior of each controller. A *controller constraint* specifies the behaviors of the controller that must be satisfied to avoid UCAs. In general, each UCA can be inverted to define constraints for each controller.

Step 4: Identify Loss Scenarios

Once unsafe control actions have been identified, the next step is to identify loss scenarios, as shown in Figure 2.8 [Leveson and Thomas, 2018].

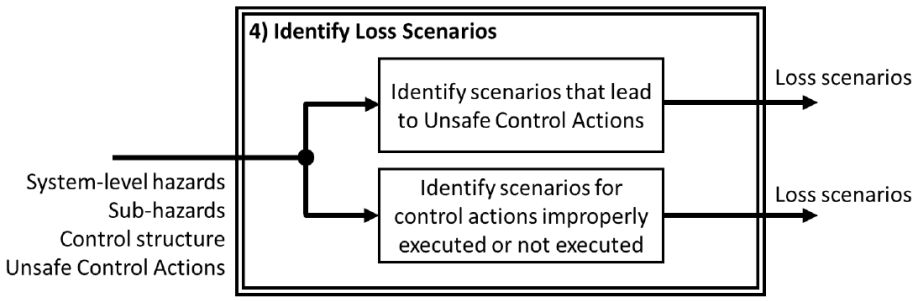


Figure 2.8: Overview of the fourth stage of STPA: identification of loss scenarios [Leveson and Thomas, 2018].

A loss scenario describes the causal factors that can lead to unsafe control actions and hazards. Two types of loss scenarios should be considered: a) Why would Unsafe Control Actions occur? and b) Why would control actions be performed inadequately or not performed at all, leading to hazards?

At this point in the analysis, as illustrated in Figure 2.9, two components must be added: actuators and sensors, which implement control actions and *feedbacks*, respectively.

Identifying scenarios that lead to UCAs

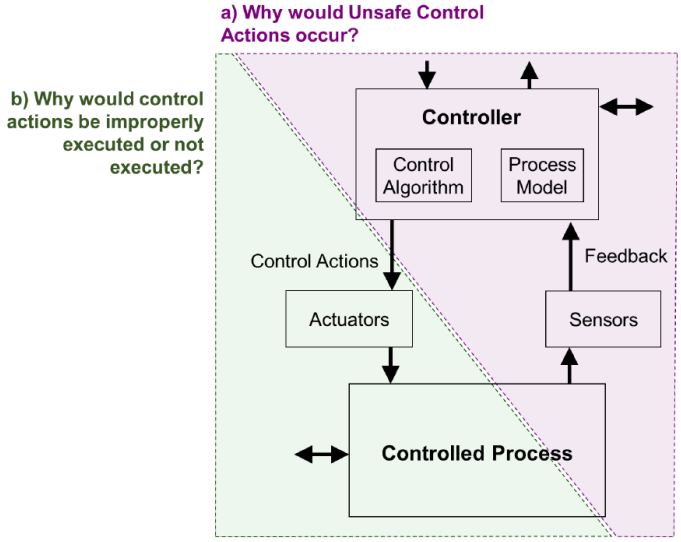


Figure 2.9: Identify loss scenarios [Leveson and Thomas, 2018].

Such scenarios can be constructed by starting with a UCA and reasoning backward to identify the factors that might cause the controller to provide (or fail to provide) that control action. In general, scenarios leading to UCAs may include:

- Controller-related failures (for physical controllers):
  - Physical failure of the controller itself;
  - Power failure;
  - Others.
- Inadequate control algorithm:
  - Faulty implementation of the specified control algorithm;
  - The specified control algorithm is flawed;
  - The specified control algorithm becomes inadequate over time due to changes or degradation.
- Unsafe control input:
  - UCA received from another controller (already addressed when considering UCAs from other controllers).
- Inadequate process model:
  - The controller receives incorrect *feedback*/information;
  - The controller receives correct feedback/information but misinterprets or ignores it;
  - The controller does not receive feedback/information when needed (delayed or never received);
  - The necessary feedback/information from the controller does not exist.

Identifying scenarios in which control actions are performed incorrectly or not performed at all

Hazards can be caused by UCAs, but they can also be caused without a UCA if control actions are performed incorrectly or not performed at all. To create these scenarios, we must consider factors that affect the control path as well as factors that affect the controlled process.

2.4 Agile Project Planning

The purpose of project planning is to create and coordinate effective plans that define and understand the objectives, assumptions, constraints, and scope of a project [ISO/IEC/IEEE, 2023]. Planning is not inherently part of the (traditional) RE process, but their tasks are closely intertwined [Nawrocki et al., 2014]: both activities require the project scope to be defined. The connection between planning and RE becomes even more evident in agile development contexts, where planning and continuous project deliveries align with the main practices of the RE process [Cao and Ramesh, 2008].

Agile approaches discard long periods of preliminary analysis but recognize the importance of defining the project's directions and understanding its requirements [Caroli, 2018]. In this context, a challenge often faced in this practice is finding the balance between agility and efficiency, focusing on artifacts that add value to the project [Inayat et al., 2015, Menezes et al., 2021]. Project planning and RE in the context of agile development share this concern.

Planning is a critical step for the success of a project [Zwikael et al., 2014], and has been the subject of interest in several studies with different perspectives, ranging from the proposal of models and tools [Balfe et al., 2017] to the level of stakeholder involvement [Heravi et al., 2015]. However, many organizations face difficulties in incorporating project planning practices, making it a challenge, especially for those that deal with complex projects and need to involve various stakeholders [Rosacker and Rosacker, 2010].

Whether in project planning or initial RE activities, the goal should not be to produce a highly detailed plan or comprehensive requirements specification, as these can quickly become outdated. Instead, the focus should be on creating artifacts that guide the project in the right direction and serve as inputs for subsequent stages. In this context, the canvas approach has proven effective in defining the scope and capturing essential information (e.g. from a project) especially when collaboration among stakeholders is critical, increasing clarity and facilitating the analysis and communication of ideas.

2.4.1 Using Canvas in Planning Activities

A canvas is an artifact for prototyping a mental and visual model applicable to the design and planning of projects, businesses, or other purposes. As a strategic planning tool, its primary objective is to address fundamental questions related to the object of analysis. Each of these questions is represented by components (elements that encapsulate and detail essential information according to the type of canvas), forming an interconnected structure that describes the intended subject. The canvas follows a logical structure to construct a visual map, helping to organize and refine ideas, and should be accessible, visible, and collaboratively adaptable as needed.

In general, a canvas is created and employed with the purpose of supporting the planning of specific activities. The Business Model Canvas (BMC), proposed by [Osterwalder and Pigneur, 2010], for example, is one of the most traditional and widely adopted canvas models, serving both as a powerful tool for business planning and as a reference for other canvas types. Beyond the BMC, several other models have been proposed in the literature to support activities such as project planning and other purposes [Thiée, 2021].

2.4.2 Reference Models

Next, we present the two main canvas models that were adopted as the basis for the development of our proposal, which we refer to as *reference models*. We emphasize that the other canvas models mentioned throughout this chapter fall within the same set of characteristics as the reference models, and are often proposed based on them.

Business Model Canvas (BMC)

As described above, the canvas approach has been increasingly used in planning activities and in various fields of application. Analyzing this scenario, [Osterwalder and Pigneur, 2010] pioneered the development of a canvas to plan and describe, through a simplified model, a logic of creation, delivery, and capture of value for a given business. The Business Model Canvas (BMC) comprises new components, according to the model[3] presented in Figure 2.10, and has a wide scope in the area of business planning. The success and effectiveness of the BMC have established it as a reference model, which has been widely adopted and used as a basis for the development of other types and models of canvas for various fields of application.

[3] https://www.strategyzer.com/library/the-business-model-canvas

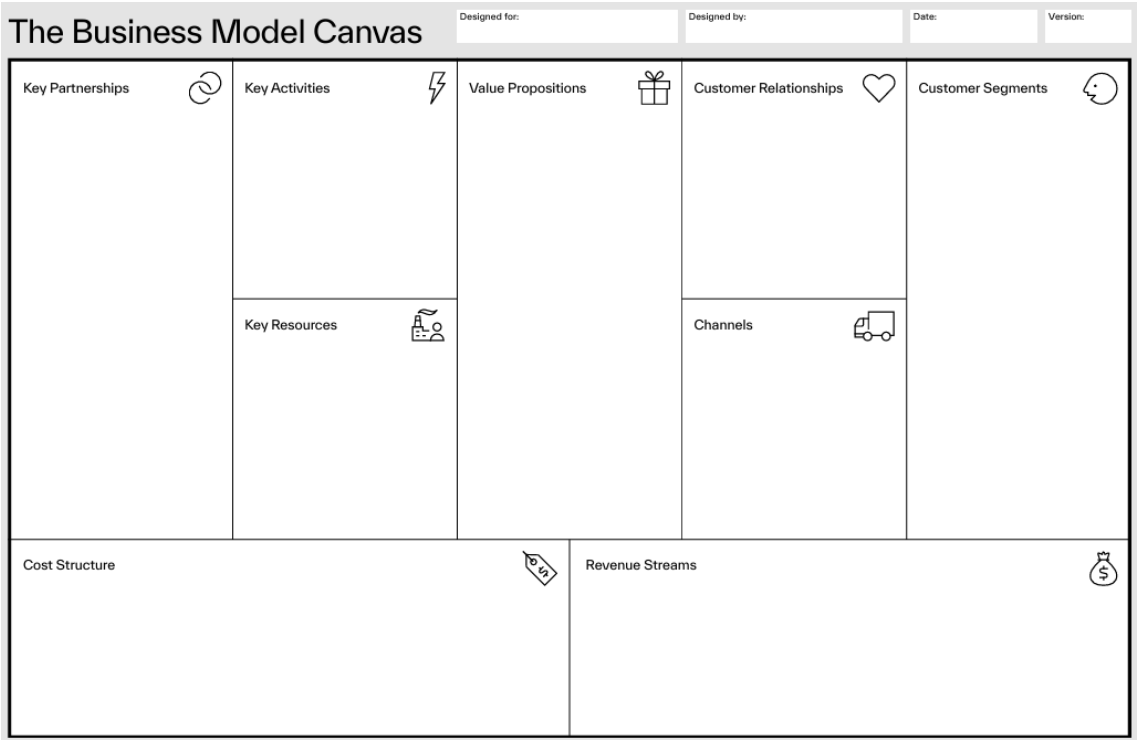


Figure 2.10: Template of the *Business Model Canvas* BMC.

Project Model Canvas (PMC)

Exploring a more comprehensive and efficient project planning model, Finocchio [Finocchio-Júnior, 2013] developed the *Project Model Canvas* (PMC). While BMC focuses on business design, PMC offers an agile and efficient approach to project planning, incorporating 13 key components for this purpose, as shown in Figure 2.11. These components are based on project management concepts, which are interconnected and complement each other in a logical sequence for completion and validation.

2.4.3 Methodological Support for Canvas

Currently, the use of canvas is widespread, meeting various needs and applications in several fields of application. However, after conducting an ad hoc review of the different types of canvas, covering a series of primary studies, for example, [Caroli, 2018, Silva et al., 2021, Takeuchi et al., 2022] and a secondary study [Thiée, 2021] (systematic review), we highlight the lack of methodological support for the construction of these artifacts, which implies problems such as: (i) lack of standardization; (ii) inconsistencies between models proposed for the same domain; (iii) low comprehensibility of the elements of a canvas model; (iv) difficulties in reusing or extending reference models; (v) misuse of comprehensive models for specific application domains; and (vi) low effectiveness of models due to one or more precedent cases.

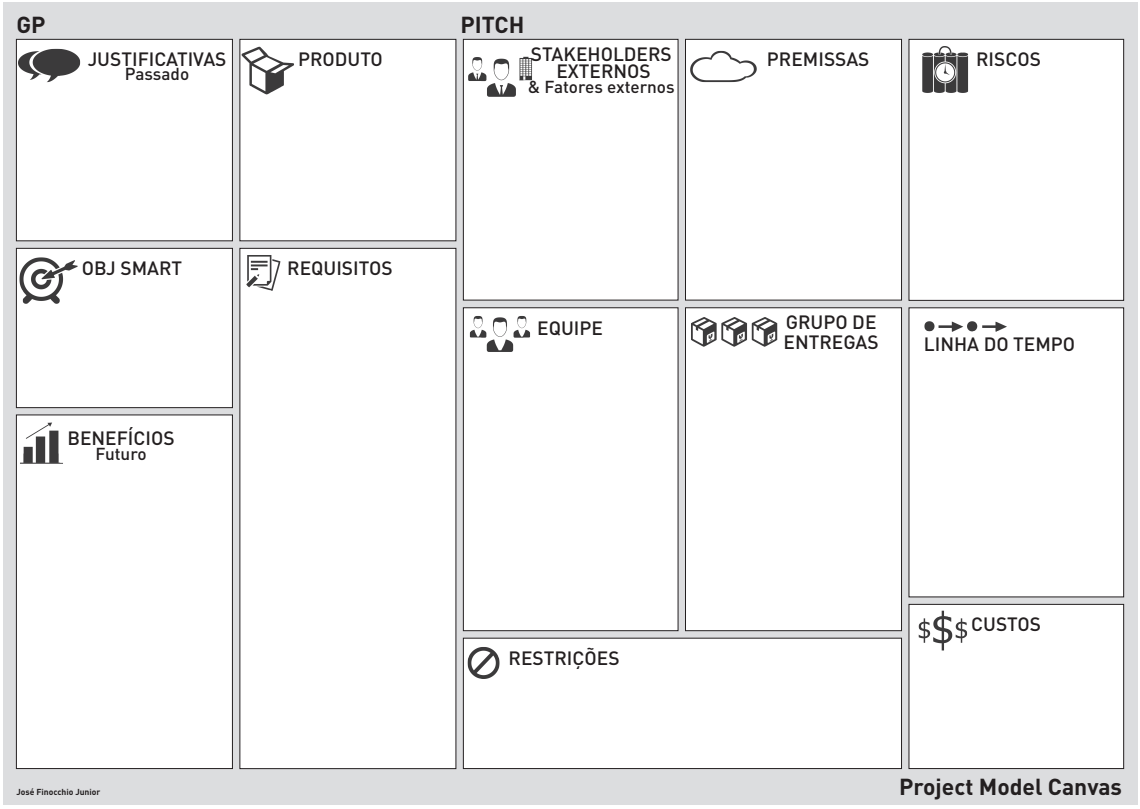


Figure 2.11: Template do *Project Model Canvas* BMC.

To summarize this set of problems identified in the literature, we will refer to them as *lack of methodological support for canvas*. In order to address and tackle these problems and provide the necessary methodological support, we propose, as part of the contributions of this research, a metamodel to support the development of (models of) canvas. The abstraction process of this metamodel is presented below.

2.5 Chapter Summary

Chapter 2 establishes the essential conceptual basis for this research, beginning with the definition of IoT systems and their close convergence with CPS, understood as the integration of digital resources and connectivity with physical devices and systems. The central focus of the work is on critical IoT systems. For such systems, it is crucial to address safety (related to accidental failures) and security (related to intentional threats) in an integrated way, since a weakness in one dimension can compromise the other. RE for both properties is typically risk-oriented, emphasizing the specification of behaviors that must not occur.

In light of the limitations of traditional safety analysis approaches (such as FMEA and FTA) in handling the complexity and interdependence of modern systems, the chapter presents STPA. It is a top-down risk analysis method grounded in systems theory,

which conceives safety as an emergent property and focuses on control and feedback structures rather than isolated component failures. This approach is particularly suitable for complex systems such as IoT and can be extended to the unified analysis of safety and security. The STPA process unfolds in four main steps: defining the purpose of the analysis (including losses and hazards), modeling the control structure, identifying UCAs, and identifying loss scenarios.

The chapter concludes with a discussion of agile project planning and the use of visual artifacts such as canvas models. Planning is a crucial activity closely connected to RE, as it supports scope definition and the initial elicitation of requirements. Reference models such as the Business Model Canvas (BMC) and the Project Model Canvas (PMC) are highlighted for their effectiveness as collaborative tools for structuring complex information and aligning stakeholders. However, a lack of methodological support for constructing these models is identified. Altogether, these theoretical foundations (addressing the nature of critical IoT systems, integrated risk analysis via STPA, and project structuring through canvas) form the backbone for the proposal and construction of the thesis artifacts presented in the following chapters.

CHAPTER 3

Related Work

This chapter presents an analysis of studies related to the proposal developed in this thesis. These studies were identified through a literature review process carried out since the beginning of the research, initially through a Systematic Literature Mapping (SLM) and, subsequently, a Systematic Literature Review (SLR) that deepened the most relevant results of the mapping. The purpose of conducting systematic studies was twofold: to ensure reproducibility and to establish well-defined protocols that document and organize the research process.

Section 3.1 outlines the scope and objectives of the SLM, along with its main findings, which highlighted the need for an additional review stage focused on more specific objectives: safety and security requirements. Section 3.2 provides an overview of the SLR results and the studies identified, emphasizing how safety and security co-analysis has been addressed in complex and critical systems. Section 3.3 discusses the main research gaps identified, while Section 3.4 presents the final considerations.

3.1 Requirements Engineering for IoT/CPS Systems: Systematic Literature Mapping

SLM is a form of evidence synthesis that follows the structure of a protocol defining research questions, research strategies, selection criteria, and others. This type of study allows the identification of evidence clusters and research gaps that direct the focus for further research [Kitchenham and Charters, 2007].

At the beginning of this doctoral work, we comprehensively investigated the RE process for IoT/CPS systems, based on initial studies in the area [Wiesner et al., 2015][Wiesner et al., 2017][Garro et al., 2019]. These studies discuss the growing complexity of these systems, which has led to greater effort required from designers and engineers in terms of understanding and solving the problems encountered. To this end, an SLM on the topic was conducted in 2021, the preliminary results of which can be found in [Veiga and Bulcão-Neto, 2022]. The results of this publication

were subsequently extended as more relevant works were analyzed. Below, we present an overview of this SLM.

3.1.1 Objective, research questions, and search strategy

The objective of the SLM[1] was to comprehensively understand the RE process for IoT/CPS systems. To guide this study toward the proposed objective, four main research questions were defined:

RQ1. What trends have been identified in the state of the art in RE for IoT/CPS systems?
RQ2. How are RE practices applied in IoT/CPS systems in the literature?
RQ3. What are the main contributions of RE studies for IoT/CPS systems?
RQ4. What are the main challenges and open questions highlighted in RE for IoT/CPS systems?

An automatic search strategy was adopted[2], validated by an RE researcher, in the digital databases ACM DL, Engineering Village, IEEE Xplore, and Scopus. The search string, presented below, was defined after a series of tests and adjustments for refinement (according to protocol), with the assistance of an experienced RE researcher and other systematic studies in the field [Lim et al., 2018][Kaleem et al., 2019][Kudo et al., 2020].

```
("requirements engineering" OR "requirements analysis" OR "requirements
specification" OR "requirements model*" OR "requirements elicit*" OR "requirements
management" OR "requirements gather*" OR "requirements validation" OR "requirements
collect*" OR "requirements identif*" OR "requirements documentation" OR
"requirements verification")
AND ( "IoT" OR "internet of things" OR "cyber-physical systems")
```

3.1.2 Analysis of results

For reasons of space, the details of the study selection and data extraction stages are reported in [Veiga and Bulcão-Neto, 2022], which consolidated the preliminary results of the SLM. That publication analyzed 31 studies relevant to the objectives of the review, published between 2020 and 2021. Subsequently, works from 2019 were incorporated, bringing the total to 49 relevant studies identified and analyzed in the SLM.

In [Veiga and Bulcão-Neto, 2022], the treatment of safety and security requirements was identified as one of the main challenges and open issues in RE research for IoT/CPS systems. The importance of this topic is underscored by the fact that 24 of the

[1]The complete SLM protocol can be found at: https://bit.ly/MSLProtocolo.
[2]Conducted at all bases and completed on February, 2022.

49 studies addressed safety and/or security requirements, with 14 of them focusing specifically on RE approaches dedicated to this type of requirement, given their critical role in such systems.

An in-depth analysis of these 14 studies identified important gaps to be investigated in the context of safety and security requirements, such as: the need for joint treatment of these requirements in critical IoT systems, the interdependence between security and safety, problems related to conflict identification and resolution, and the complexity of the RE process for these requirements. In order to investigate and understand these gaps, it was found that a new study focusing on security and safety requirements for critical IoT systems was needed. The objectives and results of this study are presented in Section 3.2.

3.2 Safety and Security Requirements for Critical IoT Systems: Systematic Literature Review

With the aim of furthering research into safety and security requirements for critical IoT systems, an SLR was conducted. This review took as its starting point 14 works focusing on safety and security relating to SLM. From this initial set of studies, the snowballing forward technique was applied to identify all articles that cited these initial works, in order to verify the evolution of research that has investigated safety and security requirements in critical IoT systems.

The SLR search phase returned 142 studies, of which 36 were initially selected as relevant. The remaining studies were excluded because they met one of the exclusion criteria: i) it is not a primary study, ii) it does not address safety or security requirements, iii) it does not apply to IoT systems, and iv) the text is not available. Of the 36 studies analyzed during the extraction stage, 23 were included as relevant for data extraction and analysis.

The results of the analysis show that the focus of the identified approaches is on safety/security analysis activities. In addition, analysis approaches and techniques for safety and security requirements for IoT systems have evolved to adapt to the growing complexity and criticality of these systems, replacing traditional techniques.

3.2.1 Paradigm shifts in safety and security

While traditional safety approaches worked well for simple systems in the past, significant changes have occurred in the systems being developed today and also in the context in which they are developed. Several changes are pushing the boundaries of safety engineering [Leveson, 2016]:

- **Rapid pace of technological change**: technologies are changing faster than engineering techniques can respond to these changes; as a result, new technologies can introduce unknown factors into systems and create new avenues for losses; lessons learned about design and accident prevention may be lost or become ineffective when old technologies are replaced by new ones, which has been happening at an increasingly rapid pace.
- **Changes in the nature of accidents**: as technologies and society change over time, so do the causes of accidents; Systems engineering and safety engineering techniques have not kept pace with technological innovations, and many of the approaches to predicting accidents that worked for electromechanical components are ineffective for accidents arising from the use of digital systems and software.
- **New types of hazards**: Advances in society are creating new types of hazards, exposing large numbers of people to risky situations; the most common safety strategies have limited impact on many of these new hazards.
- **Increased complexity and greater coupling**: complexity is increasing in different ways, especially with regard to the integration between the components of a system; the operation of some systems is complex and challenges understanding, which is limited to specialists, and sometimes information about the systems is incomplete; complexity makes it difficult for system designers and operators to deal with all situations, especially those related to safety.
- **Reduced accident tolerance**: accidents are increasingly expensive and catastrophic, often causing irreversible damage; learning about accidents or losses is increasingly necessary to predict new occurrences.
- **Difficulty in selecting priorities and tradeoffs**: while potential losses in accidents are increasing, companies are dealing with aggressive and competitive environments in which cost and productivity become the rule; there is greater pressure to create shortcuts and prioritize cost and schedule risks than to ensure safety.
- **More complex relationships between humans and automation (socio-technical systems)**: increasingly, humans are sharing control of systems with automation and taking high-level decision-making positions with automation implementing those decisions; these changes lead to new types of human error; inadequate communication between humans and machines is becoming a major factor in accidents, and traditional safety approaches are not prepared to address these types of errors.
- **Regulatory changes and public perception of safety**: in the complex and interrelated structure of today's society, responsibility for safety is shifting from the individual to the government; ways are needed to develop more efficient regulatory strategies without hindering economic objectives.

Traditional approaches to safety analysis, designed for systems that did not share the concerns mentioned above, are not effective for analyzing increasingly complex systems [Yu et al., 2021]. For example, a modern vehicle is a system consisting not only of tens of thousands of physical components, but also of large amounts of software code. However, most traditional approaches begin by decomposing the system and analyzing the components independently, which leads to neglecting the impacts of interactions between components. Furthermore, traditional causality models attribute accidents to an initial component failure cascading through a set of other components [Young and Leveson, 2014] and fail to address the causes of losses with nonlinear cause-and-effect links.

In this sense, new and more powerful safety analysis techniques, based on systems theory, are being developed and used successfully in a wide variety of systems today, including aircraft, nuclear power plants, automobiles, medical devices, and so on. In addition, many of the safety changes analyzed by [Leveson, 2016] also apply to security engineering, and the systems theory on which established techniques such as STAMP/STPA are based can similarly provide a powerful foundation for security analysis. An additional benefit of these approaches, which has been extensively investigated by the work presented in this chapter, is the potential to create an integrated approach to safety and security, as initially proposed by [Young and Leveson, 2014].

Current IoT/CPS systems strongly integrate physical processes and information and communication technologies. As today's critical infrastructures are often composed of complex systems, ensuring their safety and security becomes of paramount importance. In this sense, traditional safety analysis methods, such as HAZOP, have proven inadequate for evaluating these systems [Friedberg et al., 2017]. Furthermore, cybersecurity vulnerabilities are not considered critical in some systems because their effects on physical processes are not fully understood.

Security analysis is an essential activity for identifying potential system vulnerabilities and security requirements in the early stages of design. Due to the increasing complexity of modern systems, traditional approaches are inefficient in identifying unsafe incidents caused by complex interactions between physical systems, human and social entities [Yu et al., 2021]. On the other hand, STPA-based approaches can analyze losses as results of interactions between components, focusing on controlling system vulnerabilities rather than identifying external threats, and are applicable to complex socio-technical systems.

In these complex and critical systems, safety and security are two essential properties. Therefore, it is imperative to identify the interdependencies between safety and security in order to comprehensively assess and manage potential risks. In this sense, a systematic framework with concrete implementation steps is needed to perform

a combined co-analysis of safety and security [Glomsrud and Xie, 2019]. The STPA approach was developed for safety analysis of systems that can be represented using well-defined control system designs. However, the STPA approach was initially developed with a focus on security and, for this reason, has been the subject of study and improvement to enable adequate analysis of today's complex systems, which, in addition to security concerns, are often critical in terms of security.

In this sense, the SLR conducted identified a series of approaches based on STPA, which perform extensions and adaptations to fill gaps related to co-analysis of safety and security. Approaches that integrate STPA with other techniques were also identified in order to meet specific needs, mainly related to security, creating approaches referred to here as hybrid. The main safety and security approaches identified in the SLR are presented and discussed in Subsections 3.2.2 and 3.2.3.

3.2.2 Approaches based on STAMP/STPA

This subsection presents security analysis approaches based on extensions or adaptations of the STAMP approach or the STPA method.

STPA-SafeSec

The study by [Friedberg et al., 2017] presents STPA-SafeSec, a new analysis methodology for safety and security based on STPA. Its results show the dependencies between cybersecurity vulnerabilities and system safety. Using this information, the most effective mitigation strategies to ensure system safety and security can be readily identified. The study applies STPA-SafeSec to a use case in the power grid domain and highlights its benefits.

The premise of this study is based on the STPA approach to security STPA-Sec [Young and Leveson, 2013], which suggests that security analysis is always performed to ensure system safety. Furthermore, this study indicates that there is a difference between using STPA for safety and STPA-Sec for security. However, safety and security need to be addressed collectively, and according to [Friedberg et al., 2017], this collective approach is possible with STPA.

However, considering that only a collective analysis of safety and security can allow the analyst to identify the dependencies between both properties and obtain the most optimized results, the STPA presentations existing at the time of publication of [Friedberg et al., 2017] have a set of limiting factors:

- There is currently no guidance for an integrated approach to safety and security using STPA in which both are treated as equally important and mutually influencing system properties. In its original formulation, STPA-Sec argues that safety is only

relevant insofar as it affects security. This represents a limited and reductionist view of the system. In the socio-technical context of modern cyber-physical systems, monetary loss to the operator should also be considered a critical loss. Such losses may arise, for example, from breaches of confidentiality (e.g., consumer data or intellectual property) without direct implications for safety. Consequently, traditional STPA-Sec must be extended to allow analysts to account for losses that are not directly related to safety.

- The STPA-Sec approach also lacks guidance on how to conduct security analysis once the critical aspects of the system have been identified. First, it does not extend the causal factors defined for the safety domain to the security domain, which makes analysis more difficult and reduces comparability between results of different studies. Moreover, STPA-Sec does not provide mechanisms to integrate well-established security analysis techniques, thereby limiting its applicability in comprehensive system assessments.

STPA-SafeSec aims to address the shortcomings mentioned above and introduce the following improvements over traditional STPA methods.

1. The description and evaluation of a unified approach to safety and security analysis based on STPA and STPA-Sec. This approach equally prioritizes safety and security and allows for the detection of a broader set of hazard scenarios.
2. An extension of the causal factor orientation, which is focused on safety in STPA, to the security domain. This extension should provide support for the analyst and make the results more comparable.
3. A method for linking the abstract control structure to the physical system design to integrate the results of traditional security analysis methods. Based on the physical system design, traditional security analysis techniques can be used in a manner complementary to STPA-SafeSec.

Although it provides improvements related to safety and security co-analysis, the approach does not address specific relationships between safety and security requirements, nor does it address strategies for identifying and resolving conflicts.

STPA-Extension

The study by [Glomsrud and Xie, 2019] explores the feasibility of applying STPA-based safety and security co-analysis to new CPS. The authors identify a gap between the first two steps of STPA, considering the difficulties of modeling control structures (in Step 2) based on the systems engineering foundation established in Step 1.

In this sense, the first contribution of the study is to fill this gap by extending the analysis performed in Step 1. More specifically, functional requirements, derived

from safety (or security) constraints, are used to facilitate the modeling of control structures. The STPA safety and security co-analysis was studied to assess the risks of the CPS collectively. The second contribution of the study was, according to the authors, to improve the STPA-Sec [Young and Leveson, 2013], appropriately integrating widely adopted security analysis methods. Thus, a comprehensive list of vulnerabilities and threats can be identified through the improved analysis process, and specific security requirements can be derived.

When applying conventional STPA or STPA-Sec approaches to the analysis of emerging CPSs, such as autonomous ships, several challenges arise [Glomsrud and Xie, 2019]. First, Step 1 does not provide sufficient analysis artifacts to support the design of new systems, which are often still in the exploration phase and therefore not well defined. Likewise, modeling the control structure in Step 2 is difficult when prior knowledge and practical experience with such systems are limited. To address these challenges, a structured approach for STPA-based safety and security co-analysis has been proposed.

The STPA-Extension broadens the definition of safety to include losses related to availability, security, and efficiency, thereby extending the scope of the analysis. In addition, the authors highlight a gap between the constraints identified in Phase 1 and the control structure defined in Phase 2. To bridge this gap, the approach proposes translating the identified constraints into functional (safety and security) requirements, which can then guide the development of the control structure. The improvements introduced in Phase 1 also influence the co-analysis carried out in Phase 4, where not only accidental and unintentional causes are examined at the component level, but also intentional ones. Overall, two distinct aspects must be considered independently in the co-analysis process: (i) security-related causes that may lead to safety-related losses, and (ii) safety-related causes that may lead to security-related losses.

STPA-SynSS

Currently, security and safety attempt to prevent losses in complex software-controlled systems. As discussed earlier, the STPA approach focuses on identifying unacceptable consequences related to security failures. In this context, STPA-SysSS [Zhou et al., 2021] is a new STPA-based methodology that proposes to synthesize safety and security to perform the complete hazard analysis process.

The hazard analysis process is a formal process that includes hazard identification, assessment, and control. Based on the identification of hazard effects and causal factors, STPA-SynSS can be used to determine the importance of hazards and establish design measures that will eliminate or mitigate the identified hazards. As a continuous iterative closed-loop process, STPA-SynSS can effectively track potential hazards until an

acceptable process closure action is broken and verified. Furthermore, in this approach, cybersecurity is considered a strategic issue rather than a tactical one, and therefore STPA-SynSS focuses on how to control system vulnerabilities rather than how to avoid threats. The proposed redesign of the co-analysis method makes the analysis of security against attacks broader in relation to the design of a system. These changes not only prevent system losses caused by unknown threats, but also prevent system losses introduced by unknown threats, such as malicious individuals.

The objective of the proposed method is to address the specifications identified in existing alternatives and introduce the following improvements.

1. STPA-SynSS frames the entire STPA-based risk analysis process and elevates safety and security to equal priority. Hazard elimination/mitigation strategies are implemented in the system design through system safety and security requirements, so that hazards can be continuously tracked and closed-loop management of potential hazards can be implemented.
2. The impact of security on system safety is considered from a high-level "strategic" perspective rather than a "tactical" perspective. The method avoids not only disruptions from known threats, but also disruptions introduced by unknown threats, such as potential system intruders. The main issue faced in preventing security-related losses was focused on "unintended" design that creates system vulnerabilities.
3. Security concerns were included in unacceptable losses. In addition, a distinction was made between the definitions of the terms "accident" and "loss," where an accident/incident can lead to one or more losses. Intermediate steps for identifying accidents (safety) and incidents (security) were integrated into the analysis process.
4. To overcome the limitation of prior knowledge required to build the functional control structure, functional requirements derived from safety/security constraints were introduced.
5. Generate the mapping between abstract control loops and physical components to refine constraints and derive more specific causal factors, which shows the visualization of the identified model elements and their responsibilities.
6. The generation of loss scenarios was extended to the domain of security, as a starting point for further reflection and investigation of the intentional and malicious activation of unsafe control actions. The identification of loss scenarios was differentiated into two distinct categories: inadvertent scenarios and intentional scenarios.

The study presented matures some key factors in the process of co-analysis of safety and security, inserting new inputs into the process to enable improvements. However, the proposed method does not comprehensively address the interactions between safety and security, nor does it concern itself with identifying and resolving conflicts.

Cybersafety

The study by [Khan and Madnick, 2022] presents a holistic and integrated analysis of safety and security, called Cybersafety, which is based on the STAMP framework, for an industrial cooling system as an example scenario. In this analysis, vulnerabilities emerging from the interactions between technology, operator actions, and organizational structure are identified, and recommendations are provided to systematically mitigate the resulting loss scenarios.

The Cybersafety method proposes additions to STAMP to further focus on the analysis of security issues and uncover systemic vulnerabilities:

- In Step 1, system-level hazards should be formulated as system-level threats. This implies reframing hazards in terms of attacker targets. This would help bridge the terminology gap with traditional security analyses by focusing the analyst's attention on "thinking like an attacker."
- In Step 2, controllers that impose security controls should be explicitly identified, in addition to controllers that impose security restrictions. Furthermore, for each controller, security responsibilities should also be explicitly identified alongside safety responsibilities.
- In Step 3, *insecure* control actions (from a security standpoint) for each controller must be identified in conjunction with *unsafe* control actions (from a safety standpoint).
- In Step 4, loss scenarios should cover not only the application of safety constraints, but also security constraints.

The study reports filling a gap in the current application of STAMP-based theoretical systems methods for security (such as STPA-Sec) in that these approaches were essentially security analyses that identified causal factors for cyber security. The authors argue that security should be considered an integral part of the safety control framework, with explicit identification of security-related roles and responsibilities, constraints, control actions, and loss scenarios. To this end, the proposed improvements to the Cybersecurity method (focusing on threats, security constraints, and controllers) provide a well-guided and structured approach to developing an integrated safety and security model to holistically identify cyber vulnerabilities and mitigation requirements in complex industrial control systems.

Despite its advantages, the Cybersecurity method does not provide any indication of the prioritization of loss scenarios, unlike risk-based methods, which provide a risk score. In this sense, future work is cited to improve the method for prioritizing loss scenarios so that those responsible for security can address the most critical weaknesses in their systems supported by an order of priority.

3.2.3 Hybrid approaches

This subsection presents approaches based on a combination of methods aimed at improving the security and safety analysis process.

STPA-DFSec

The work of [Yu et al., 2021] seeks to fill the gaps in STPA-Sec [Young and Leveson, 2013] related to the lack of attention to non-safety-related security issues, such as information security (e.g., data confidentiality) and the lack of efficient guidance for identifying concepts in this domain. To this end, [Yu et al., 2021] propose an adaptation of STPA-Sec based on data flow (called STPA-DFSec) to overcome the aforementioned limitations and systematically elicit security constraints.

The authors use STPA-DFSec and STPA-Sec to analyze a vehicle digital key system and investigate the relationship and differences between both approaches, their applicability, and highlights. The conclusion of the study by [Yu et al., 2021] is that the proposed approach can identify information-related problems more directly, due to the data processing aspect. As an adaptation of STPA-Sec, it can be used with other STPA-based approaches for the co-design of multidisciplinary systems under the unified STPA framework.

In summary, the STPA-DFSec approach follows the general structure of STPA but introduces a data flow-based structure for information security considerations. Thus, STPA-DFSec reorients the scope of the STPA-based security analysis approach to consider more information security-related issues that are not covered by STPA-Sec. To achieve this adaptation, it was necessary to add features and modify some steps in relation to STPA-Sec, allowing for efficient guidance to better support this information security analysis based on the STPA framework.

Two limitations of STPA-DFSec were identified. The first is a general limitation of STPA-based approaches, which lack evaluation of the identified scenarios. By evaluating and ranking the identified loss scenarios, the system designer can decide which unsafe scenario should be avoided with high priority. To overcome this limitation, STPA-based approaches can be used in combination with other metrics to evaluate the identified unsafe behaviors and scenarios. The second limitation is that the analyst needs to have the corresponding information about the data processing of the target system (e.g., how data flows between components and what kind of data processing function is contained), otherwise the functional interaction structure cannot be constructed.

SafeSec Tropos

The growing convergence of information technology with operational technology and the corresponding proliferation of interconnected CPSs has given rise to several safety and security challenges. One such challenge relates to the systematic identification of coherent, consistent, and non-conflicting safety and security requirements. The study by [Kavallieratos et al., 2020b] proposes an integrated method for engineering safety and security requirements for CPSs at the design stage of the system lifecycle. The method identifies safety and security objectives, systematically elicits a comprehensive list of requirements, and links these requirements to the objectives, thereby facilitating the conflict resolution process.

Safety and security objectives describe the system features that ensure the safety and security of the system. An essential step in the proposed method is the identification of these objectives for each system under analysis. Constraints related to safety/security objectives can influence the safety and security analysis and design of a CPS. A safety/security dependency introduces safety/security constraint(s) that must be respected in order for the corresponding dependency to be satisfied.

The proposed method is based on the integration of two methods, Secure Tropos and STPA, from which safety and security requirements are elicited by analyzing system safety and security constraints.

3.3 Research gaps identified

Based on the analysis, gaps were identified in the following areas:

- **Lack of support tools to guide joint safety and security analysis**:
  - This gap is evident in the first phase of the STPA process, especially for the design of new systems, which can create a lack between defining the purpose of the analysis and modeling the system control structure, as identified by [Glomsrud and Xie, 2019][Carreras Guzman et al., 2021];
  - The remaining stages of the co-analysis process need to reflect changes in the definition of the purpose of the analysis in a systematic and well-founded manner. Some studies have sought hybrid approaches, which have proven interesting for addressing joint safety and security analysis;
  - This factor can lead to a larger and unnecessary number of iterations in the process or to incompleteness and loss of requirements in the process. In addition, a more detailed process through the definition of specific artifacts for safety and security co-analysis could reduce the complexity of the analysis and

even support its implementation by software engineers who are not specialists in safety and security.

- **Treatment of interdependencies and conflict resolution:**
  - The proposed approaches do not perform a comprehensive analysis of the relationships between safety and security in terms of conditional dependency, mutual reinforcement, and antagonism between the results of safety and security analysis [Sadvandi et al., 2012][Kriaa et al., 2015];
  - The identification and resolution of conflicts between safety and security is explicitly addressed in only one of the approaches analyzed, referring to the study by [Kavallieratos et al., 2020b], showing that there is a greater need for research on this topic in safety and security requirements.
- **Documentation of the analysis in safety and security requirements specifications:**
  - The studies analyzed did not address techniques for documenting and communicating the results obtained to relevant stakeholders, a factor highlighted as a gap in approaches to safety and security co-analysis studies [Kavallieratos et al., 2020a];
  - STPA-based methods generally produce a series of artifacts, however, they are focused on specific stages of the analysis; an artifact focused on specification could improve the documentation of results in terms of validation and management of requirements, providing an interface between safety analysis and other stages of system development.
- **Lack of tools to support the analysis processes and proposed approaches:**
  - Only one of the methods presented was directly supported by a software tool for its implementation.
  - Although STAMP/STPA-based tools exist, they are limited to the scope of these methods.
  - The lack of tools to support the proposed approaches is a factor that may limit their use, since safety and security co-analysis methods are applied to complex systems where software support is of great importance [Kavallieratos et al., 2020a].
  - The development of tools to support this process is cited as future work by [Yu et al., 2021][Zhou et al., 2021].

3.4 Chapter Summary

This chapter presented a summary of the results of two systematic studies of the literature that investigated, respectively, the RE process of IoT/CPS systems and approaches that address safety and security requirements for critical IoT systems. The first study identified the research concern with safety/security requirements in the context of IoT systems, and how the relationship between these requirements can impact the design and development of these systems. Based on this premise, the second study was conducted using the snowballing forward technique on the relevant works from the first study, which allowed the identification of a series of works that investigate the STPA-based co-analysis of safety and security.

The chapter also discusses the main advances and limitations of existing approaches, highlighting the need for more robust methods for co-analysis of safety and security in critical systems and identifying research opportunities, such as the development of artifacts that promote greater integration between safety and security, the creation of tools to support the analysis and specification process, and improvements in the documentation and communication of the results obtained.

Based on the results of these systematic studies, works related to the proposal of this doctoral research were identified, as well as important research gaps that require efforts to advance the state of the art regarding the safety and security RE process in critical IoT systems.

CHAPTER 4

MM4Canvas: A Metamodel for Methodological Support and Canvas Construction

This chapter presents the process of designing and developing a metamodel for canvas: *MM4Canvas*. This metamodel was proposed to offer methodological support and assistance in building canvas models, which has been identified as a gap in the literature. The motivation for devising a canvas metamodel in the context of this research was the need to systematically establish an artifact to support agile project planning that could be used at the beginning of the RE process for critical IoT systems, considering: i) the identification of the scope and essential components of these systems and ii) as support for the elicitation of their safety and security requirements.

4.1 Abstracting a Metamodel for Canvas

In this section, we present the process of abstracting a metamodel for methodological support and assistance in building canvas-based models. The canvas approach has been used in the literature to meet different types of planning demands, but without methodological support for the development of these models. In this sense, based on the literature, we identified the need and opportunity to establish an abstraction for this type of artifact.

4.1.1 Metamodeling Approach for Canvas Abstraction

Model-Driven Engineering (MDE) is an approach to software and systems development that prioritizes models as the primary artifacts throughout the development process. MDE posits that these models are not merely descriptive or documentary, but are actively involved in the development, maintenance, and operation of systems. This paradigm shift places a strong emphasis on abstract representations (of artifacts), allowing developers to work at a higher level (of abstraction) to represent various aspects of a

system. In this way, models serve as abstractions to capture essential characteristics, support reasoning, analysis, and transformation.

Models are constructed using Domain-Specific Modeling Languages (DSMLs), defined by a metamodel [López-Fernández et al., 2015]. DSMLs offer specialized languages for specific domains, allowing the creation of models that capture essential concepts and relationships. A metamodel is essentially "*a model that defines the structure of a modeling language*" [Rodrigues da Silva, 2015], providing a formal description of the syntax, semantics, and constraints of the language. Thus, the *MM4Canvas* metamodel is an effort of this research to define the structure of a modeling language for canvas, in support of the development of these types of models.

Metamodels describe the types of entities in a domain and the ways in which they can be associated, providing a vocabulary and rules for building models. Thus, a metamodel can be used as a formal and structured basis for building and managing artifacts throughout the life cycle of a software engineering project. This approach contributes to the definition of rules, standards, and structures that guide the creation and organization of artifacts, bringing consistency and quality to the process [Henderson-Sellers, 2011]. The metamodeling approach can contribute in several ways to the development of artifacts:

- *Definition of common structures*: metamodeling establishes a common vocabulary and structure for artifacts such as models, diagrams, requirements documents, and design specifications. This makes it possible to create artifacts that follow a specific standard, facilitating understanding and integration between different teams and stakeholders involved in a project.
- Reusable models: through a metamodel, it is possible to create artifacts that can be reused in different types of projects, saving time and effort in developing new models.
- *Domain-specific extension*: using the metamodeling approach, it is possible to create artifacts specific to a domain by incorporating new properties and requirements into existing (more comprehensive) models, based on the elements supported by the metamodel.
- *Artifact customization*: metamodeling allows artifacts to be adjusted and customized according to the specific needs of the project or organization, promoting greater adherence to business and technological requirements.
- *Ease of updating*: using a well-defined metamodel simplifies artifact maintenance, as artifacts are structured to adapt to changes and evolutions in the system or project.

By providing a high-level structure, a metamodel helps control and standardize the development and evolution of artifacts, promoting quality, consistency, and ease of maintenance. For this reason, after systematically examining a series of canvas models

and abstracting their essential and common elements, we identified the opportunity to abstract an artifact to support canvas development. The artifact was proposed as a metamodel, with the aim of providing the methodological support necessary to instantiate, reuse, and extend various types of canvas models, and meet the specific needs of organizations and project teams.

4.1.2 Essential Elements of a Canvas

Based on the study conducted during the research to survey and analyze the various types of canvases available in the literature, we identified common elements among these models, which became the basis for the proposed metamodel. Taking the PMC and BMC, reference models presented earlier, as examples, we introduce the main elements of a canvas: fundamental questions, building blocks, and relationships.

The BMC and PMC, like other canvases identified in the literature, are developed based on building blocks, which can be defined as a category of information essential to the purpose of the canvas, which must be identified in order to achieve its objective (the planning of an activity). These building blocks have relationships between them, which should support the construction of some type of strategic planning, and can be grouped into fundamental questions, which would be broader categories of organization, presenting different perspectives on the object of analysis (the business or the project, in the case of BMC and PMC, respectively).

These fundamental questions are generally based on the idea of an action plan (5W2H), a set of questions used to compose strategic plans quickly and efficiently [Kerzner, 2002]. Table 4.1 compares the BMC and PMC models, showing the fundamental questions used by each type of canvas and the building blocks involved and integrated in each question.

Table 4.1: Fundamental questions and building blocks of canvas models: BMC and PMC.

Fundamental Questions	**Canvas Model and Type**	
	PMC (project-oriented)	**BMC (business-oriented)**
Why	Justifications, Objectives, Benefits	–
What	Product, Requirements	Value Proposition
Who	Stakeholders, Team	Customer Relationships, Customer Segment, Channels
How	Assumptions, Delivery Groups, Constraints	Key Partnerships, Key Activities, Key Resources
When	Timeline	–
How much	Risks, Costs	Cost Structure, Revenue Sources

Thus, based on a detailed analysis of the structure of different canvas models, we carried out a study to abstract, identify, and generalize the essential elements common to all types of canvas. We exemplify this abstraction through reference models (BMC and PMC), which are consolidated in the literature and used for different purposes (business and projects, respectively).

These identified elements are essential for building a canvas and common to the different models analyzed. All canvas models have building blocks that aim to support the extraction of specific information directly related to the purpose of the analysis performed. These blocks, in turn, can be grouped into fundamental questions that need to be answered during canvas-based planning. In addition, these blocks may be related to each other. These relationships are constructed in order to generate added value from the joint analysis of the building blocks involved, and can be used for bidirectional validation of the information identified when filling out a template instantiated from the model.

4.2 Metamodel for Canvas (*MM4Canvas*)

Based on studies conducted to identify elements common to different types of canvas models and in response to the challenges presented in subsection 2.4.3, we propose a metamodel to support the development and methodological support of canvas models: *MM4Canvas*. The objective of this metamodel is to provide a solid foundation for modeling, analyzing, building, and maintaining canvas models, promoting consistent standardization, reuse, and specialization in different application domains. The metamodel allows the creation of canvas models for different purposes and needs, including project planning (which is the focus of this research), business planning, among others.

4.2.1 *MM4Canvas* Requirements

Below we describe the main requirements identified for the development of the *MM4Canvas* metamodel:

- Generalization: the metamodel must allow the instantiation of different types of canvas models for different purposes and application domains;
- *Standardization*: the metamodel must allow different canvas models to be instantiated from the same types of elements, having a standardized structure and facilitating the extension and reuse of its building blocks;
- *Extensibility*: the metamodel should allow and facilitate the extension of general-purpose canvas models, supporting the creation of new models that serve specific application domains;

- *Reusability*: the metamodel should allow and facilitate the reuse of building blocks for the development of other canvas models;
- *Traceability*: the metamodel should allow the construction of relationships between elements and the definition of a traceability process based on links, from the most specialized information or element to the most general.

4.2.2 Methodology for the Development of *MM4Canvas*

As discussed in the presentation of the research methodology, we adopted Design Science Research (DSR) due to its problem-solving nature and systematic creation of artifacts. In this research, we instantiated and applied the DSR-Model [Pimentel et al., 2020], which has three main pillars:

- The *problem in context* (based on the state of the art): the lack of methodological support in the development of canvas models;
- The *behavioral conjectures* (based on the theoretical foundation of the research): the metamodeling approach contributes to the systematization of the canvas development process, establishes standardization, and supports the reuse and extension of models;
- The *proposed artifact* (aimed at addressing the state of practice): its design is guided by the conjectures and addresses the problem): a metamodel for canvas, which we call *MM4Canvas*.

By analyzing different types of canvas that have been adopted in the state of the art [Osterwalder and Pigneur, 2010][Finocchio-Júnior, 2013][Caroli, 2018], we identified the essential elements present in all canvas models (fundamental questions, building blocks, etc.) and the different types of relationships between them (inheritance, composition, association, etc.). By modeling these elements and their relationships, we propose a metamodel that allows the instantiation of different types of canvas models for various purposes, from the most generic to the most specialized, aiming to meet different planning needs and domains.

4.2.3 Metamodeling Architecture

The proposed metamodel, called *MM4Canvas*, adopts a widely used metamodeling architecture: the *MetaObjectFacility* (MOF) [OMG, 2002]. In this architecture, the elements of the lower layers are instances of those in the immediately higher layers, as shown in Figure 4.1, described below:

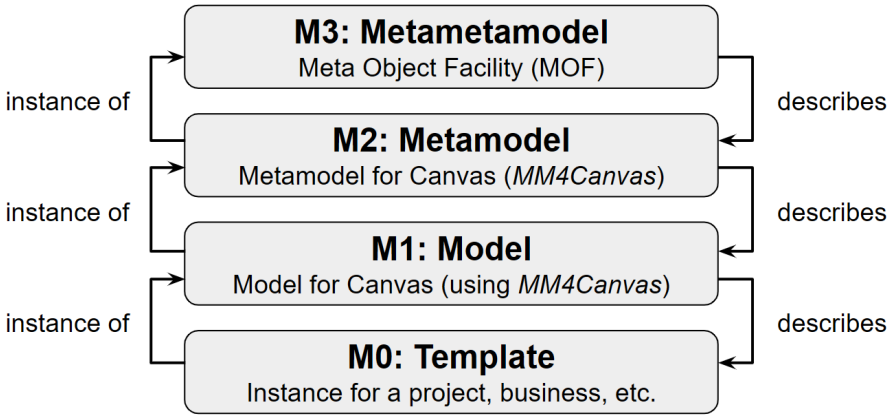


Figure 4.1: Integration between MOF and *MM4Canvas*.

- *Layer M3* (meta-metamodel): this upper layer includes the MOF itself, which defines the meta-metamodel used to create other metamodels. It essentially provides the schema for defining how metamodels are structured, thus eliminating the need for a higher level of abstraction.
- *Layer M2* (metamodel): this layer contains the metamodels that define the structure and semantics of models relevant to specific domains.
  - Example: the *Unified Modeling Language* (UML) resides at this level, offering a standardized notation (class diagrams, sequence diagrams, etc.) that can be used in different software development projects to model different aspects of the system, such as architecture and behavior.
  - The proposed *MM4Canvas* metamodel also resides at this level, providing the necessary elements for the development of canvas models for different application domains.
- *Layer M1* (model): this layer contains models that describe specific instances (whether of systems, software processes, or other types of abstractions/artifacts) according to the rules and structures defined in layer M2. These models are direct instances of the metamodels and are customized to meet the specific needs of the business/application.
  - In the context of this research, models are representations of different types of canvas, which can be developed and customized for different application domains. These models can be used both for the implementation of software that allows the use (filling and manipulation) of the canvas from outside dynamically, and for the creation of a canvas template that can be used physically or digitally.

- *Layer M0* (template): this layer contains the actual data or runtime instances that the models in layer M1 represent. This is where the developed models are put into practical use, either within a software environment or through a template.
  - In the context of this research, instances are canvas templates for different purposes, filled with planning information according to the building blocks of the type of model represented.

Thus, the proposed metamodel, *MM4Canvas* (layer M2) is an instance of MOF (layer M3) and describes the elements of possible terminal models, which are the different types of canvas models that can be instantiated from the metamodel. These canvas models (layer M1) are the basis for creating canvas software/documents/templates (i.e., artifacts) (layer M0) for projects, businesses, or other purposes.

4.2.4 *MM4Canvas* Constructs

The *MM4Canvas* metamodel, shown in Figure 4.2, was established based on a set of elements (constructs) essential for abstraction and conception of canvas models, detailed below:

- The ***canvas*** (metaclass *Canvas*): this is the basis for constructing a canvas artifact. It is a composition of fundamental questions, which in turn are a composition of building blocks that together form a canvas model. It may contain attributes or metadata that describe essential information about the instantiated model (name, date, version, etc.).
- The ***fundamental questions*** (metaclass *FundamentalQuestion*): these are high-level questions based on the idea of an action plan (5W2H) that articulate essential aspects of the project. A fundamental question consists of one or more building blocks. They offer different perspectives on the project by addressing key questions such as “what” needs to be done, “who” will do it, etc.
- The ***building blocks*** (metaclass *BuildingBlock*): these are the elements that encapsulate the essential information or specific requirements for planning the project, business, etc. They are described through one or more items, according to the need for detail of the objective. They can be specialized in:
  - *General-purpose elements* (metaclass *GeneralPurposeElement*): these are building blocks that describe more comprehensive information about the objective. In the case of a project-oriented canvas, for example, whose reference model is PMC, general-purpose blocks are based on classic project management concepts (as shown in Table 4.1). These building blocks are grouped

into fundamental questions according to the type of information they describe about the project.

  - *Domain-specific elements* (metaclass *DomainSpecificElement*): these are building blocks that describe information or requirements specific to a domain. They can be added and used as a mechanism for extending reference models (e.g., PMC or BMC) or as a way to modify/specialize general-purpose elements to expand the descriptive capacity of a canvas model, aiming to serve projects or businesses (among other objectives) with specific needs or requirements. It must be associated with a specific domain (metaclass *Domain)*.

- The ***items*** (metaclass *Item*): information that describes each building block according to its purpose for planning based on the canvas type. Each building block can be associated with (*isDescribedBy*) one or more items that describe it.
- The ***relationships*** (metaclass *Relationship*): building blocks can be linked by relationships, which define relevant associations between them, according to the needs and objectives of the canvas model. Each relationship between these elements (metaclass *ElementRelationship*) starts "from" one building block and goes "to" at least one other, which may even be the same block.

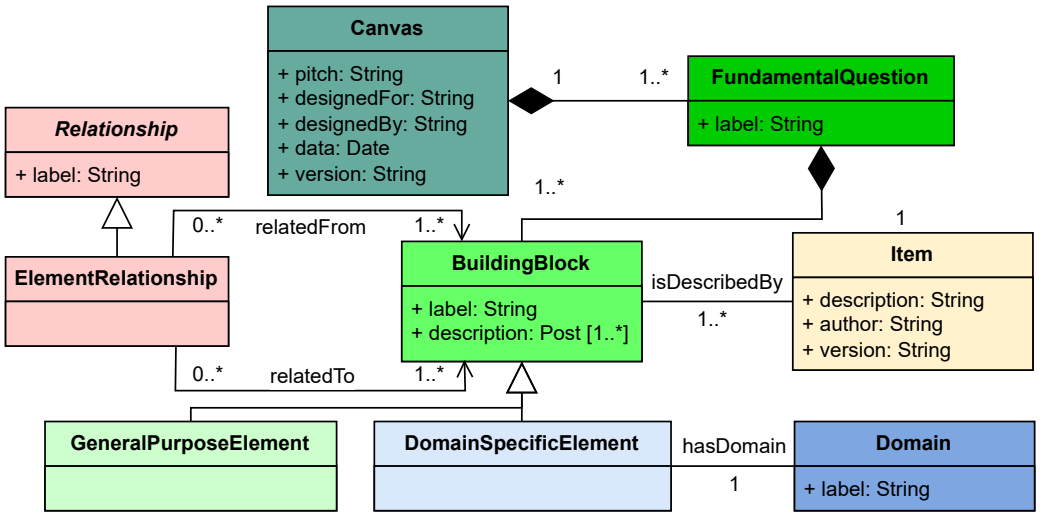


Figure 4.2: *MM4Canvas*: metamodel (M2) to support the creation of canvas models.

Because it is based on essential elements of canvas construction, rather than elements or relationships focused on specific objectives or domains, the *MM4Canvas* metamodel provides methodological support and can be used to develop canvas models for different purposes (projects, businesses, etc.) and various application domains or needs (IoT systems, security, etc.). Thus, a canvas model instantiated from *MM4Canvas* comprises a set of fundamental questions defined and instantiated based on the objectives

of the canvas model being constructed, aiming to support strategic planning according to its purpose.

4.2.5 Granularity and Traceability of Elements

This subsection presents the levels of granularity present in the context of the elements of the proposed metamodel and the traceability relationships between models and instances, created from *MM4Canvas*.

Granularity of the elements of the *MM4Canvas* Metamodel

Granularity is a term used in different contexts, usually related to the level of detail or division of something. Depending on the area of study or application, granularity can take on specific meanings. In Software Engineering, for example, it refers to the level of detail or division of components in a system. Granularity is considered i) fine when the elements are divided into smaller and more specific parts, or ii) coarse when the elements are more aggregated or generalized.

Bringing the concept of granularity to the context of the *MM4Canvas* metamodel, we can analyze it by considering different types of granularity when observing its elements. Granularity changes as we observe the canvas from each of its elements, from the most comprehensive to the most specific, as shown in Figure 4.3 and explained below:

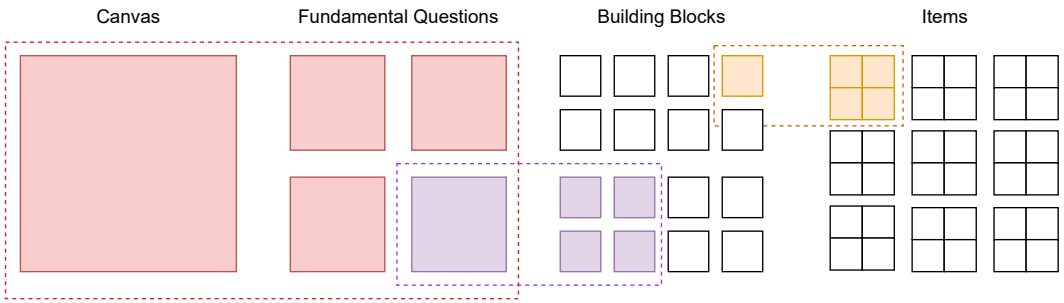


Figure 4.3: Granularity level of the elements of the *MM4Canvas* metamodel.

- The canvas is the most comprehensive element, as it is a composition of fundamental questions;
- Each fundamental question is composed of a set of building blocks, increasing the level of detail of the elements of a canvas;
- Each building block is described by a set of items, which are the finest granularity elements, and therefore the most specific, of the metamodel.

Traceability between elements and instances of *MM4Canvas*

Traceability is defined as "the ability to relate artifacts created during the development of a software system to describe: i) the system from different perspectives and levels of abstraction, ii) the stakeholders who contributed to the creation of the artifacts, and iii) the rationale that explains the form of the artifacts" [Spanoudakis and Zisman, 2005]. Traceability also refers, in the context of requirements, to the ability to define, capture, and follow the traces left by requirements in other elements of the software development process (artifacts) and the traces left by these elements in the requirements [Pinheiro, 2004].

The perspectives of granularity and traceability of the elements of the metamodel are important both for understanding the structure of the canvas models that will be instantiated and for constructing the relationships between the elements: whether these are between elements based on the metamodel itself or with elements of other related artifacts that will be used in the RE process. Traceability is the characteristic (and requirement) of the metamodel that will allow, as it is instantiated, to track the origin, relationship, and impact of one element to another over time and activities performed, such as elicitation, analysis, and specification.

Analyzing granularity at the building block level, it is at this level that the internal relationships (between blocks) of a canvas model are constructed. These relationships contribute to increasing the expressiveness of the model and even to the subsequent validation of planning information. Furthermore, considering this same level of granularity, there may be traceability between the building blocks of a canvas model and elements of other artifacts, building a source-destination relationship between activities in a process, for example, initiated by the planning activity based on a canvas model.

Granularity at the item level, after a model has been instantiated and contains planning information, allows this information to be directly mapped or used in the following stages of an RE process, for example. In addition, traceability at this level is essential for identifying changes and impacts on the scope of planning. If a change is necessary in an earlier activity in the process, and this directly impacts essential planning information, traceability at the item level will allow the description information of a canvas element that needs to be updated to be identified.

Thus, granularity and traceability are intertwined in the design of the proposed metamodel: granularity defines the level of detail of the elements that can be traced, both internally (within the same artifact) and externally (between different artifacts), which are defined as vertical and horizontal traceability [Pinheiro, 2004], respectively. Finer granularity (at the item level) improves the traceability of planning information. Therefore, understanding the appropriate level of granularity at which one is working is essential to define and also meet specific traceability needs.

4.3 Proof of Concept: Instantiating a Canvas Model from *MM4Canvas*

To validate the proposed *MM4Canvas* metamodel, we instantiate a canvas model (M1) for the PMC. As presented in subsection 2.4.2 and Figure 2.11, the PMC is a canvas proposed by [Finocchio-Júnior, 2013] for planning different types of projects. Due to its scope, in the context of this research, we classify it as a general-purpose, project-oriented canvas. The PMC has its building blocks based on key project management concepts, which are essential to underpin the start of a project. Next, we demonstrate how the PMC was instantiated in a canvas model based on *MM4Canvas*.

The model for the PMC, instantiated from *MM4Canvas*, is shown in Figure 4.4. All of its elements, from fundamental issues to its building blocks (general-purpose elements), were instantiated based on the definitions of the proposed metamodel and in alignment with the PMC methodology [Finocchio-Júnior, 2013]. The relationships between these elements were also mapped, defining their connections. This model (M1) serves as the basis for project planning activities and can be instantiated in software or a template to obtain answers to fundamental project questions through its project management-based elements, as well as to validate the information gathered (M0).

The first part of the model comprises the "why" of the project, where the justifications (problems to be addressed and solved), objectives, and benefits are described as elements that are intrinsically related. The justifications must consider the benefits, to which they are directly correlated, in order to solve the identified problems. The objectives should act as a bridge between the before/after scenario of the project, effectively addressing the problems described and guiding the project towards the expected benefits. These three building blocks are essential for understanding the project and why it is necessary. This is achieved by defining which problems will be addressed, the objectives, and what is desired to be achieved as results.

The second part of the model comprises "what" should be described about it and a set of essential requirements, allowing its scope to be defined. These requirements are motivated by the problems presented in the project justification and its objectives, and should refer directly to the product. In addition, other building blocks of the canvas can provide context and motivation for defining these requirements, which can be defined throughout the process of filling out the canvas.

Requirements should be defined based on quality criteria such as necessity, clarity, consistency, verification, and traceability. The use of a standard, such as ISO 29148, which defines principles and good practices for requirements, guides the elicitation, analysis, validation, and management, in addition to supporting the structuring of requirements, helping to ensure clarity, consistency, and completeness.

The building blocks corresponding to the fundamental questions “who,” “how,” and “when/how much” (addressed collectively in the PMC) are intrinsically interconnected, as illustrated in Figure 4.4. Ensuring that the information collected by each block is aligned with the defined relationships is vital to validating the accuracy of an instance of this canvas. Project stakeholders should act as “owners” of the requirements and provide assumptions for the project. The team is responsible for the delivery groups and is associated with the constraints linked to them. Project risks originate from assumptions and can threaten deliveries, requiring organization into a schedule and guiding project costs.

By using the *MM4Canvas* metamodel to map fundamental problems, building blocks, and relationships inherent to the PMC methodology, we aim to improve the effectiveness of the resulting canvas model. First, it can be used as a visual tool for project planning and initial requirements gathering by project teams and stakeholders. In addition, it supports the validation of the generated artifact. Finally, as an instance of *MM4Canvas*, this PMC model can be reused and expanded for various project purposes within specific application domains, as will be presented in Chapter 5.

4.4 Chapter Summary

This chapter presented the design and development process of the *MM4Canvas* metamodel. The proposed metamodel arose from gaps identified in the literature regarding methodological support for canvas models. Given the need for this research to plan critical IoT system projects to support the STPA-based safety and security analysis process, we saw an opportunity to contribute with a canvas abstraction that would also serve other projects and research. Thus, we propose a canvas metamodel that can serve different purposes and scenarios that require a canvas model to meet their specific strategic planning needs.

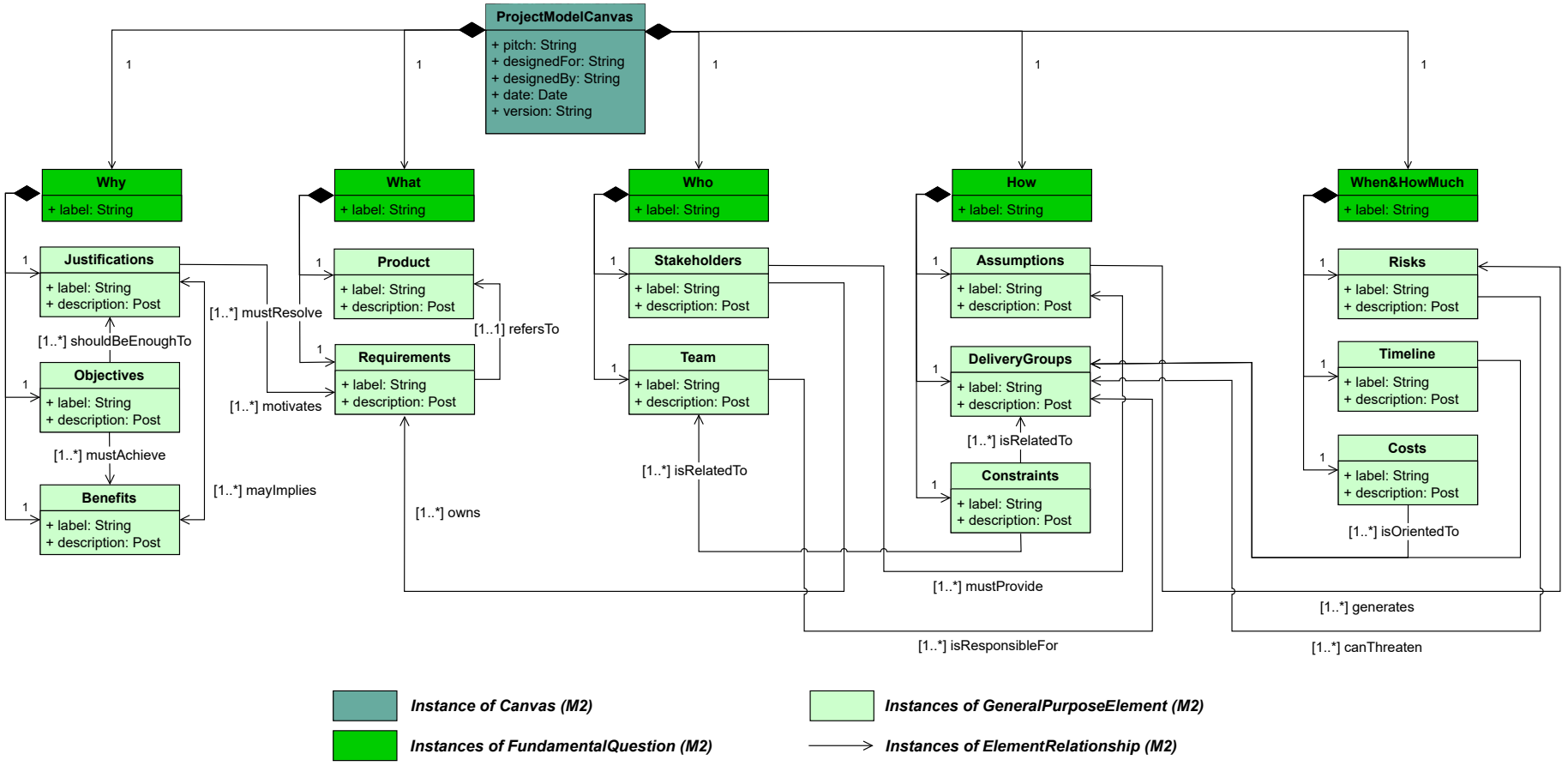


Figure 4.4: Model for PMC (M1): a canvas model for project planning, instantiated from the *MM4Canvas* metamodel (M2).

CHAPTER 5

SafeSecIoT Canvas: A Canvas Model to Support RE for Critical IoT Systems

This chapter presents the *SafeSecIoT Canvas*, a canvas model proposed to: i) support the planning of critical IoT system projects, ii) assist in the elicitation of system requirements; and iii) implement a strategy to support safety and security analysis based on STPA. The model is instantiated from the *MM4Canvas*, introduced in the previous chapter. In addition to presenting the details of the proposed model, we demonstrate how it integrates and supports artifacts for STPA-based safety and security analysis.

5.1 Project Planning and Requirements Elicitation for Critical IoT Systems

As discussed in Chapter 3, the literature presents several efforts based on the STPA method for the co-analysis of safety and security in critical systems. By starting from the definition of losses, this method identifies hazards that may result from unsafe control actions and supports the derivation of safety requirements. Approaches that extend STPA for joint safety and security analysis [Friedberg et al., 2017, Glomsrud and Xie, 2019, Zhou et al., 2021, Ribeiro and Castro, 2022, Gomola and Bouwer Utne, 2024] represent specific contributions to the analysis of critical systems. However, a gap that remains is the systematic treatment of the RE process from the very beginning of system conception, including the definition of scope, objectives, and system requirements (project planning), as well as its alignment with subsequent process activities [Glomsrud and Xie, 2019].

The RE process is a fundamental pillar upon which successful system development efforts are built. It serves as the vital bridge between stakeholder visions and the tangible product, defining the scope, objectives, and functionalities of the system [Pargaonkar, 2023]. Project planning and the RE process are closely intertwined [Nawrocki et al., 2014]. As an essential activity of project planning and RE, defin-

ing the scope and essential requirements of a system is a critical step for the success of a project [Zwikael et al., 2014], especially for those dealing with complex systems and needing to involve multiple stakeholders, as well as critical IoT systems.

Therefore, project planning and the RE process are directly linked to the overall quality of system development. A clear definition of objectives and scope helps ensure that requirements (both customer-defined and system-inherent) are properly understood, guiding development in the right direction, reducing rework, and increasing compliance with project goals [Pargaonkar, 2023]. Moreover, effective communication is essential for eliciting accurate and complete requirements, ensuring that all stakeholder needs and constraints are addressed. Hence, efficient integration between project planning and the RE process is crucial for quality assurance in the early stages of the systems development life cycle, particularly for critical systems.

However, many organizations still approach security requirements reactively, considering them only after designing and implementing the system [Mellado et al., 2010]. This practice has the potential to cause problems and introduce defects that can significantly impact the project, resulting in higher costs for correction [Jarzębowicz and Weichbroth, 2021]. Thus, the development of a security-dependent software system should address these concerns from the system design stage, in the RE phase, rather than treating it only as an aspect of the advanced stage of the development process.

In this context, after verifying the potential of the canvas approach for planning activities, and of the PMC [Finocchio-Júnior, 2013] specifically for project planning and requirements elicitation (general-purpose canvas model), we propose an extension of this model based on *MM4Canvas*: the *SafeSecIoT Canvas* (canvas model for a specific domain). The proposed canvas model aims to implement a strategy to support safety and security analysis based on STPA and reduce its complexity by including specific building blocks for extracting characteristics from systems that have these concerns.

5.1.1 Extension and Reuse based on *MM4Canvas*

To develop the *SafeSecIoT Canvas*, the building blocks (general-purpose elements) of the PMC model, focused on project management concerns, were reused, and elements specific to the IoT domain and safety and security were incorporated to meet the demands of these types of projects, extending the scope of the PMC model presented in Chapter 4. The IoT domain-specific elements defined for the *SafeSecIoT Canvas* are presented below:

- *Components* (class *Components*): refer to the hardware and software elements that make up the system. Hardware components include sensors, actuators, or any object

(thing) with identification, detection, or action behaviors and processing capabilities that can communicate and cooperate to achieve a goal. Software components include algorithms to control and orchestrate IoT systems, user interfaces, etc. These elements vary according to system requirements.

- *Connectivity* (class *Connectivity*): this is the way in which components can connect to materialize the IoT paradigm. This is not limited to the *Internet* (Wi-Fi, 4G/5G), but also encompasses technologies such as *Intranet*, *Bluetooth*, among others, that can connect components.
- *Actions* (class *Actions*): these are relevant interactions performed by the IoT system, usually between its components or subsystems, according to the context of the application, whether or not linked to actuators, and which can generate data or feedback.
- *Data* (class *Data*): any information important to the system, usually produced or originated by system components (or their interactions), which can be persisted or used in some way to support applications.

Regarding safety and security concerns, the following building blocks have been defined to *SafeSecIoT Canvas*:

- *Assets* (class *Assets*): these are all elements that have high (and critical) added value for the system. They must be protected against accidental or malicious loss. They include people, resources, the environment, or services.
- *Losses* (class *Losses*): significant damage or negative impacts associated with an asset. They are considered unacceptable to stakeholders. They can be caused by accidents or attacks.
- *Risks* (class *Risks*): these are potential causes of accidents or attacks. They may be intentional (security) or unintentional (safety), and are directly linked to the security of one or more assets.

5.1.2 The *SafeSecIoT Canvas* Model

Figure 5.1 shows the *SafeSecIoT Canvas* model (level M1), instantiated from *MM4Canvas*. The proposed model reuses the general-purpose elements of the PMC model and includes the specific elements of the IoT domain and also of safety and security, defined above.

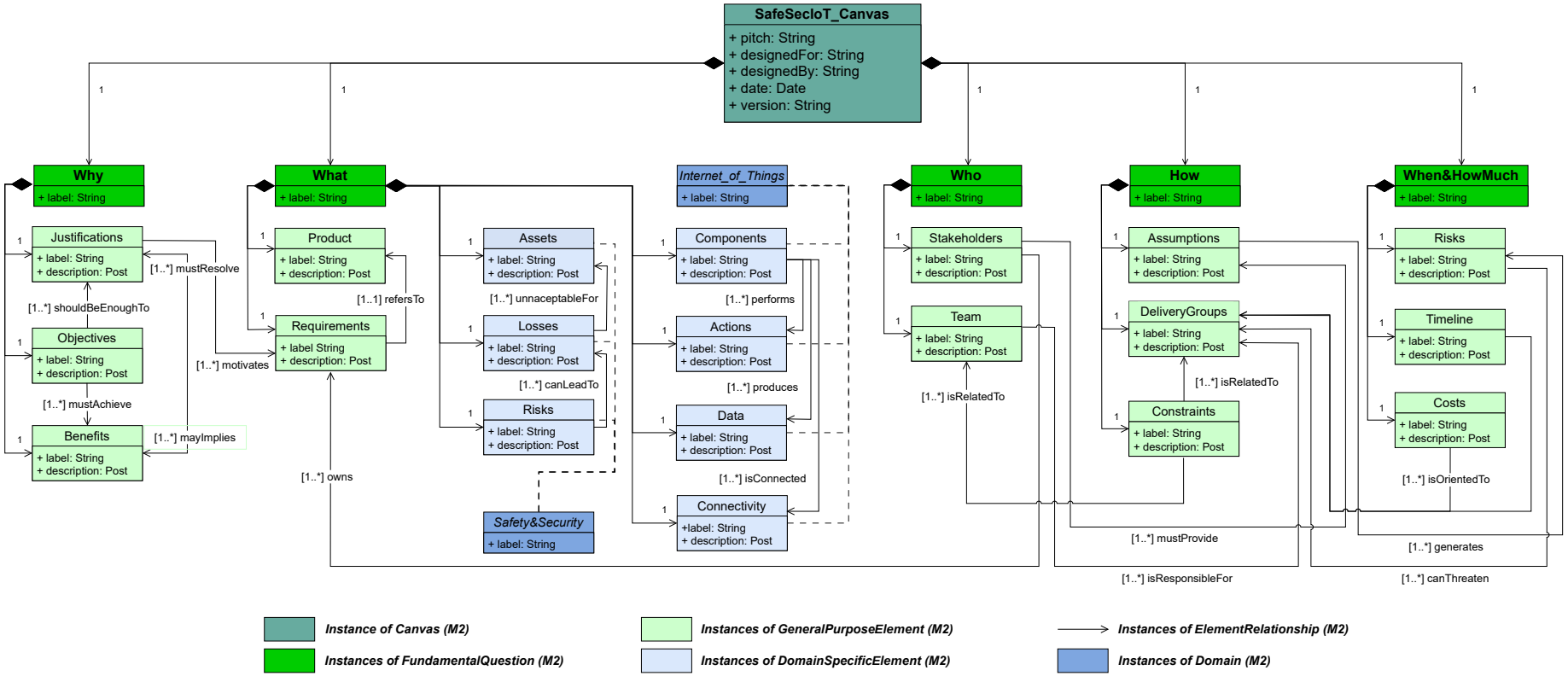


Figure 5.1: Model *SafeSecIoT Canvas* (M1): a canvas model for project planning in critical IoT systems, instantiated from the *MM4Canvas* (M2) metamodel. It includes general-purpose elements (in green, reused from the PMC model) and domain-specific elements for IoT, safety and security (in blue), thereby extending the PMC model.

Reinforcing the above, the purpose of this canvas model is to provide a tool for agile planning of critical IoT system projects, supporting the definition of essential information from the early stages of the RE process and the necessary communication between the technical team, domain experts, and stakeholders. Thus, the *SafeSecIoT Canvas* model aggregates these characteristics, based on the reuse of the PMC model building blocks, and extends it to enable the planning of critical IoT system projects, i.e., with characteristics and requirements specific to the application domain of IoT systems with safety and security requirements.

5.1.3 The *SafeSecIoT Canvas* template: instantiating the model

Considering the metamodeling architecture presented in subsection 4.2.3, a model can be instantiated using software or a template to be used at the application level (level M0) for planning activities. Therefore, a model (M1 level) can be instantiated as an M0 artifact that the responsible team and stakeholders will use in practice, within an RE process. In the context of this research, the *SafeSecIoT Canvas* model was initially instantiated in the form of a canvas template, presented in Figure 5.2, which represents the practical application of this respective model for planning the design of a critical IoT system.

Each element (general purpose or domain specific) of the *SafeSecIoT Canvas* model (level M1) represents a building block in the canvas template (level M0), which is used to extract the characteristics of the planned IoT system. The building blocks (such as objectives, benefits, actions, assets, risks, and others) are applied within the context of a critical IoT system project. The template for the *SafeSecIoT Canvas* shows how the canvas model (fundamental questions and building blocks) was instantiated and how project planning information and critical IoT system requirements can be extracted[1].

In summary, the template for *SafeSecIoT Cavas* (the M0 instance) shows how the abstract elements of the canvas model (the building blocks defined in M1) can be applied in a practical way when planning a real project. This approach ensures that the canvas model (M1) can be effectively instantiated as a template (M0), in this case, for planning a critical IoT system. Thus, the approach allows the *SafeSecIoT Canvas* artifact to be used at different levels of abstraction: i) at the M1 level, as a conceptual model to map the elements necessary for planning and their relationships; and ii) at the M0 level, as a template for practical project planning.

[1]The figure shows the template without any content, so it does not contain the items for each block. The goal is to show the representation of the *SafeSecIoT Canvas* in template format.

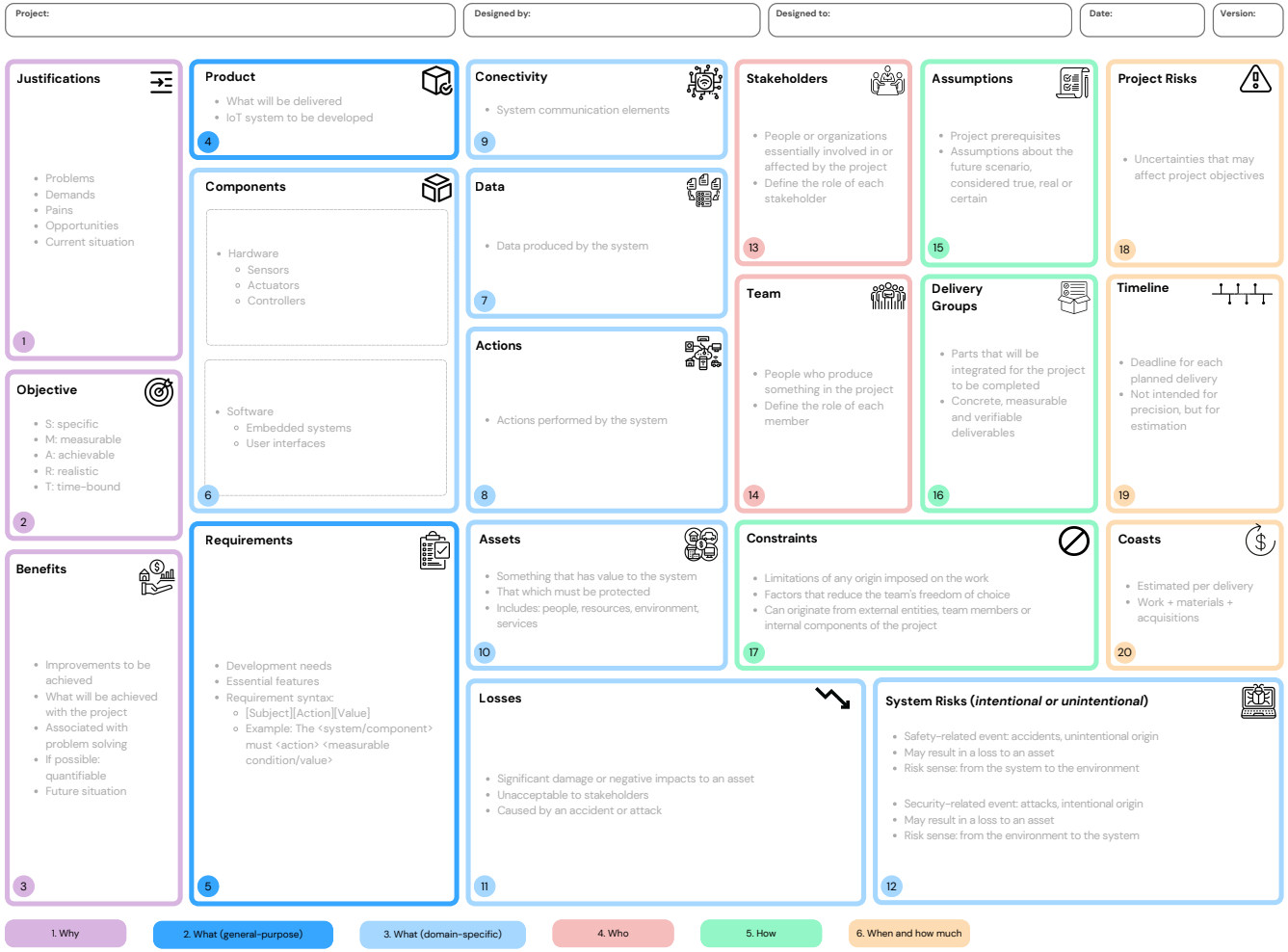


Figure 5.2: *SafeSecIoT Canvas* (M0): template for agile project planning and requirements elicitation.

5.2 Integrating Project Planning and STPA-based Analysis

This section presents the integration strategy between the artifact for planning critical IoT systems (the *SafeSecIoT Canvas*) and STPA-based security analysis methods. The essential aspects that underpin this integration proposal are also presented.

5.2.1 Adoption of ISO/IEC/IEE 15288:2023

The ISO/IEC/IEEE 15288:2023 standard [ISO/IEC/IEEE, 2023], entitled "Systems and Software Engineering - System Life Cycle Processes", is an international standard that defines a process for systems engineering throughout the life cycle of a system. It is published jointly by the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC), and the Institute of Electrical and Electronics Engineers (IEEE). The standard is widely used in different industries to ensure that complex systems are designed, developed, operated, and maintained efficiently and effectively.

The standard covers all phases of the system lifecycle and defines a set of processes and activities necessary for systems engineering, including: i) agreement processes; ii) project organizational processes; iii) technical design processes; and iv) technical support processes. This ensures that all aspects of the lifecycle are considered. In addition, ISO/IEC/IEEE 15288:2023 is designed to be compatible with and complementary to other standards, such as ISO/IEC/IEEE 12207:2021, which focuses specifically on the software life cycle. Together, these standards provide a comprehensive framework for systems and software engineering.

Below are the key contributions of ISO/IEC/IEEE 15288:2023, adopted to achieve this proposal:

- *Value of the Requirements Engineering Process*: The technical standard emphasizes the importance of understanding and documenting system requirements and using these requirements to guide its design and development. It includes specific processes for requirements engineering and system design, which are important references for the proposed safety and security RE process.
- *Risk and Quality Management*: addresses and highlights the importance of risk management and quality assurance at all stages of the life cycle. It demonstrates the importance of ensuring that risks are identified and mitigated so that system quality is maintained.

- *Adaptability*: The standard is flexible and can be adapted to different systems and organizational environments. This allows it to be used in a wide range of industries and projects, including critical IoT systems.
- *Stakeholder Involvement*: It highlights the importance of stakeholder involvement throughout the system lifecycle to ensure that their needs and expectations are met.
- *Specification and Traceability*: It establishes the need for documentation and periodic reviews to ensure that the system continues to meet requirements and function as expected.

The adoption of ISO/IEC/IEEE 15288:2023 helps organizations improve the efficiency and effectiveness of their systems engineering processes, ensuring that systems are developed in a controlled manner and meet established quality and performance objectives. In this sense, we use this standard to underpin and support the safety and security RE process and the proposed artifacts, aiming at risk management and quality assurance in the activities and tasks of this process.

5.2.2 Adapting ISO/IEC/IEEE 15288:2023 to the safety and security RE process

Below we present how the ISO/IEC/IEEE 15288:2023 processes were instantiated and adapted to support the safety and security RE activities of critical IoT systems, in the context of each proposed artifact.

Planning critical IoT system projects: *SafeSecIoT Canvas*

The planning of critical IoT system projects, carried out using the *SafeSecIoT Canvas* artifact, covers the following processes of the ISO/IEC/IEEE 15288 standard (with adaptations of activities and tasks to the intended scope):

- P1. Project planning: determines the scope of the critical IoT system project. Essential activities and tasks:
  - Define the project: i) identify the project objectives, assumptions, and system requirements; ii) define the scope of the project for the critical IoT system of interest.
  - Plan the project and technical management: i) define roles and responsibilities; ii) generate and communicate a plan.
- P2. Business or mission analysis: define the strategic problem or opportunity, characterize the solution space, and determine the potential solution. Essential activities and tasks:

  - Define the problem or opportunity space: i) analyze the problems and opportunities; ii) define the mission, business, or operational problem or opportunity to be addressed by a solution.

- P3. Define stakeholder needs and requirements: provide the necessary resources for users and stakeholders. Essential activities and tasks:
  - Identify stakeholders: individuals and classes of stakeholders who are users, supplier organizations, parties responsible for external interface entities, regulatory bodies, and others who have an interest in the system solution.
  - Define stakeholder needs: identify stakeholder needs.
  - Transform stakeholder needs into stakeholder requirements: define stakeholder requirements.

In order to make the execution of the presented processes more agile, dynamic, and efficient, all activities/tasks described are performed using the *SafeSecIoT Canvas* artifact, as shown in Table 5.1.

Safety and security analysis: STPA-based approaches

The safety and security analysis of critical IoT systems, performed using STPA-based methods, uses project planning as input and covers the following processes of the ISO/IEC/IEEE 15288 standard (with adaptations of activities and tasks for the intended scope):

- Definition of system requirements: transform the stakeholders' user-oriented vision of the desired features into a technical vision of a solution that meets the user's operational needs. Activities and tasks:
  - Define safety and security requirements: identify and specify the safety and security requirements of the system.
  - Manage safety and security requirements: i) obtain explicit agreement on requirements; ii) maintain traceability of safety and security requirements.
- System analysis: provide a rigorous basis of data and information for technical understanding to aid in decision making and technical evaluations. Activities and tasks:
  - Perform system analysis: i) apply analysis methods to identify dependencies and conflicts between safety and security requirements; ii) establish conflict resolution and conclusions or recommendations.
  - Manage system analysis: i) maintain traceability between the artifacts of the analysis process and the safety and security requirements; ii) provide the essential artifacts for the system analysis process.

Table 5.1: Relationship between ISO/IEC/IEEE 15288 processes and the *SafeSecIoT Canvas*.

ISO/IEC/IEEE 15288 Process	***SafeSecIoT Canvas***				
	Why	**What**	**Who**	**How**	**When and How Much**
Project Planning	Objective	Requirements	Stakeholders and Team	Assumptions	
Business or Mission Analysis	Justifications, Objective, Benefits	Product, Components, Actions, Data, Connectivity, Assets, Losses, Risks		Constraints	Timeline, Costs, Project Risks
Stakeholder Needs and Requirements Definition		Requirements	Stakeholders and Team		Delivery Groups

5.2.3 *SafeSecIoT Canvas* as support for STPA-based analysis: Defining the Purpose of the Analysis

Despite the intersections between STPA-based analysis and RE activities, there is a gap in the methodological coverage of system design and scope definition [Glomsrud and Xie, 2019, Veiga and Bulcão Neto, 2023]. To support the safety and security RE process for critical IoT systems from their early stages, we strategically explore the agile project planning provided by the *SafeSecIoT Canvas* to assist and streamline the definition of the system scope, providing essential inputs for STPA-based security analysis.

STPA-based analysis involves four stages with specific tasks, as detailed in subsection 2.3. In this context, the proposed strategy adopts the *SafeSecIoT Canvas* to support both Stage 1 (defining the purpose of the analysis) and Stage 2 (modeling the control structure). By strategically integrating the *SafeSecIoT Canvas* into the definition of the essential elements of the analysis, the proposed strategy directly or indirectly influences the entire STPA-based analysis process.

Figure 5.3 illustrates the strategy for integrating the *SafeSecIoT Canvas* with Step 1 (Defining the Purpose of the Analysis) of the STPA method. The strategy involves defining the boundaries of the critical IoT system using the domain-specific elements of the *SafeSecIoT Canvas*, which address IoT, safety, and security aspects. Additionally, all relevant project planning considerations are included, encompassing both general-purpose elements and other fundamental issues. This alignment between the canvas approach and STPA-based analysis aims to delineate system boundaries and enhance the definition of the analysis scope.

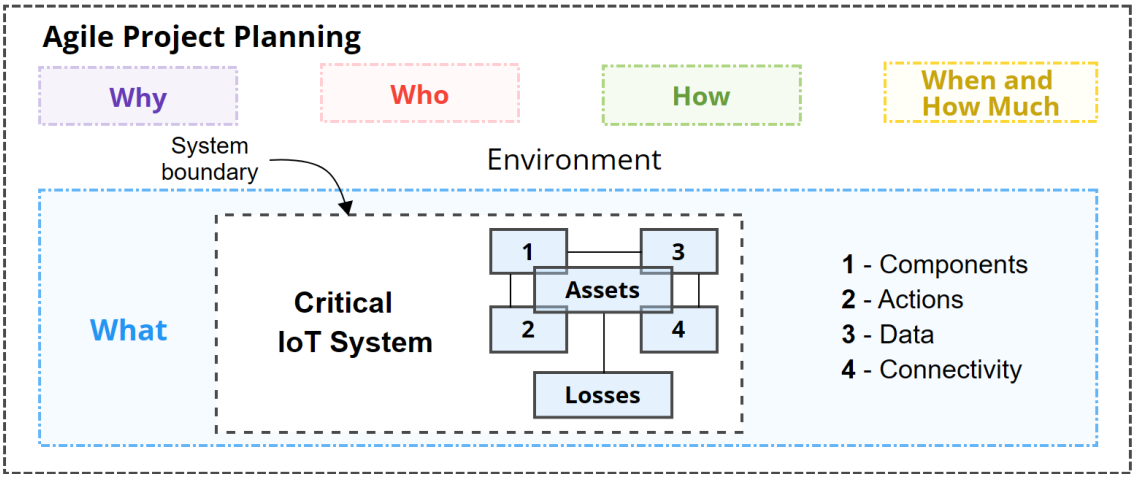


Figure 5.3: Alignment of the *SafeSecIoT Canvas* and STPA-based approaches: Step 1.

Still in Stage 1, the specific elements of the *SafeSecIoT Canvas* domain allow the definition of essential information for the safety and security analysis that will be

performed by the STPA-based approach. *Losses* are directly related to the *assets* that the system needs to protect from intentional or unintentional *risks*. In addition, the different types of losses are intrinsically linked to the relationship between assets and critical elements of the system: components, actions (mission), data, and connectivity.

5.2.4 *SafeSecIoT Canvas* as support for STPA-based analysis: Control Structure Modeling

Regarding Stage 2 of the STPA (Control Structure Modeling), the specific elements of the IoT domain of the *SafeSecIoT Canvas* were designed to relate directly to the essential elements of the system control structure and can be mapped from the planning artifact to the analysis process. Thus, the information collected through the canvas serves as input for Control Structure Modeling, as shown in Figure 5.4.

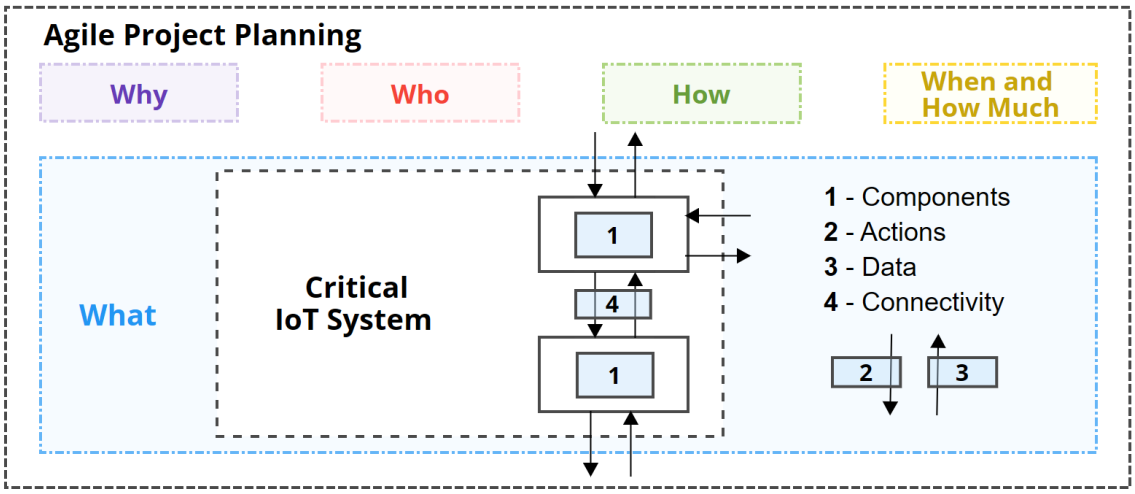


Figure 5.4: Alignment of the *SafeSecIoT Canvas* and STPA-based approaches: Step 2.

The components identified in the *SafeSecIoT Canvas* can be mapped as subsystems that will function as *controllers* or *controlled processes* within the system structure based on their responsibilities and relationships with other system components. Similarly, *actions* can be mapped as *control actions* of the system, and *data* will be mapped as *feedbacks* in the structure. *Connectivity* information will serve as the basis for understanding communication protocols between components and will be essential in identifying safety and security risks.

This strategy aims to demonstrate how canvas models can be used to support the safety and security analysis process for critical IoT systems. In addition, *MM4Canvas* allows a canvas model to be instantiated and/or adapted to the needs of different types of systems. Thus, in addition to critical IoT systems, which are the focus of this work, the proposed approach can be adopted in the context of different types of critical systems, considering their specific characteristics.

5.3 Traceability between Artifacts and in the RE Process

As discussed in subsection 4.2.5, traceability is an essential element for building the *MM4Canvas* metamodel. In addition, it allows for the development and understanding of relationships between instantiated model building blocks, as is the case with *SafeSecIoT Canvas*, as well as the integration of its elements with other artifacts of the safety and security RE process for critical IoT systems.

In subsections 5.2.3 and 5.2.4, we show how the elements of *SafeSecIoT Canvas* are used as inputs for the STPA-based safety analysis process. These mapping relationships between elements were strategically designed and developed in the proposed model so that the planning of the critical IoT system design can support and facilitate STPA-based safety and security analysis. Next, we analyze this mapping in detail and discuss the traceability of the elements within the safety and security RE process.

5.3.1 Mapping between Artifacts

Figure 5.5 shows the traceability between the elements of the *SafeSecIoT Canvas* artifact and the artifacts of an STPA-based method. We can see that there is a direct mapping between some elements of the *SafeSecIoT Canvas* and the STPA-based analysis. Based on the requirements and assets, the losses defined in the canvas artifact must be validated and supplemented, if necessary, and the risks identified in the planning must be discussed together with the losses to define hazards and threats.

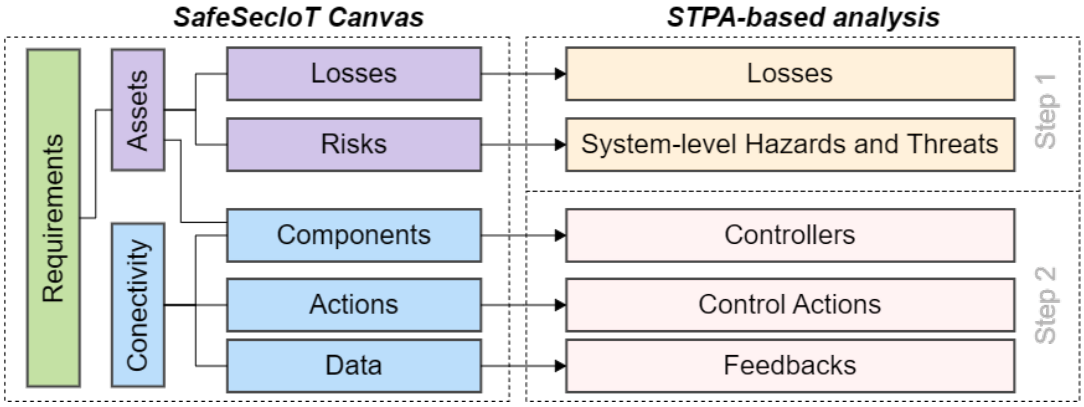


Figure 5.5: Traceability between artifact elements.

From the perspective of IoT system elements, the components identified in the *SafeSecIoT Canvas* can be mapped as controllers or controlled elements to build the control structure. Finally, the actions and data identified in the planning are mapped as control actions and feedback, completing the definition of the control structure of the analyzed system. The existing mapping between artifacts facilitates and increases the assertiveness of the STPA-based analysis process and contributes to an even more detailed

level of traceability of the analyzed elements, whose origin can be identified within the planning artifact, in a specific element for defining the scope of the system.

In addition to the direct relationships mapped in Figure 5.5, we highlight that other elements of the artifacts may be indirectly related. For example, elements specific to the IoT domain are used in the definition of *assets*, *losses* (in both artifacts), *hazards*, and *threats* in the analysis process.

5.3.2 Types of Traceability

Traceability within a model (i.e., internal traceability between elements of a single artifact) is referred to as vertical traceability, while traceability between different models or artifacts (external traceability) is known as horizontal traceability [Pinheiro, 2004]. In this proposal, vertical traceability links elements within an artifact, such as associating requirements with products (so that changes to a product are reflected in its related requirements). Horizontal traceability, in contrast, connects information across artifacts, for example, mapping items from a *SafeSecIoT Canvas* building block to inputs used in STPA-based analysis.

As presented in subsection 4.2.5, traceability is linked to the level of the elements to which it refers, that is, the granularity being considered. For example, in the *SafeSecIoT Canvas*, we can consider traceability at the level of i) artifact (*SafeSecIoT Canvas*, STPA-based method), or ii) elements (fundamental issues, building blocks, or items).

Traceability at the artifact level can be used to associate two or more high-level artifacts (e.g., the *SafeSecIoT Canvas* and the control structure of STPA-based approaches). Traceability at the element level (e.g., building blocks) can be vertical or horizontal, i.e., used internally (within the canvas to associate its building blocks) or externally (between different artifacts) to associate elements. Item-level traceability refers to specific information (or instances) related to a building block, and is therefore a finer granularity. Item-level traceability refers to planning information extracted when filling out the canvas artifact.

Figure 5.6 presents an overview of vertical and horizontal traceability perspectives based on the elements of *MM4Canvas* for planning. In the example in the figure, we refer to the planning of critical IoT system projects, where, after using the canvas model, an STPA-based analysis process is performed to identify safety and security requirements. We can see that the elements of the canvas model (building blocks) have vertical traceability between themselves (the relationships built in the model) and also horizontal traceability with artifacts from the STPA-based analysis process.

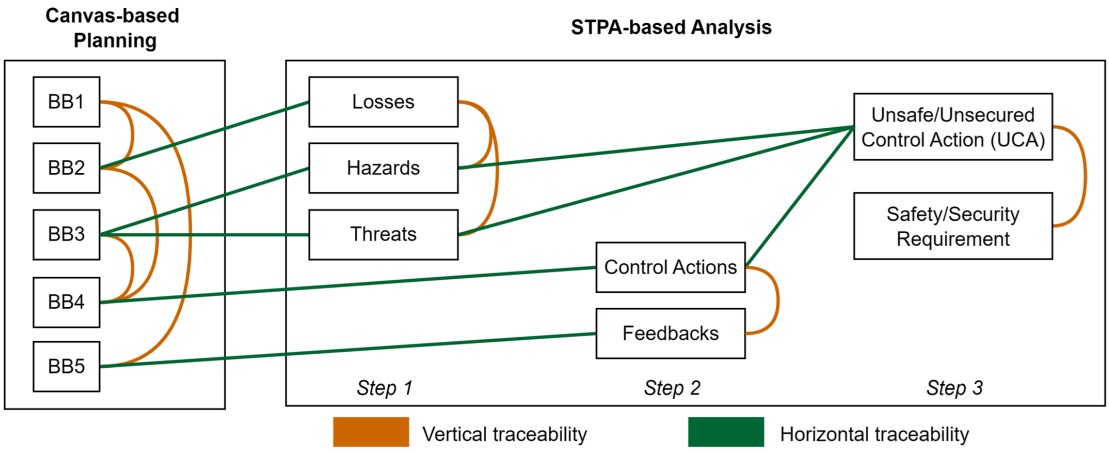


Figure 5.6: Example of vertical and horizontal traceability in canvas-based planning and other activities in the safety and security RE process.

Within the STPA-based analysis process, we can also observe vertical traceability between elements of the same stage and horizontal traceability between elements of different stages. This demonstrates that there is a traceability process that begins with project planning and continues through to the specification of requirements.

5.4 Proof of Concept: Instantiating the *SafeSecIoT Canvas* in a Critical IoT System Project

To validate the proposed *SafeSec-IoT Canvas* model for planning critical IoT system projects, we instantiated this artifact in the design of an Automated Insulin Delivery (AID) system. AID systems [Deshpande et al., 2019] close the loop between a continuous glucose monitor (CGM) and an insulin pump (IP), adjusting the insulin dose to control blood glucose in people with type 1 diabetes. The modular design of AID systems, both at the hardware and software levels, allows the use of rapidly evolving diabetes device technologies and algorithms and facilitates their use. However, this heterogeneity of components and their critical system nature also leads to growing concerns about safety and security issues.

Next, we present a proof of concept (PoC) conducted to evaluate the *SafeSecIoT Canvas* model and the proposed strategy for combining project planning and the STPA-based method in the safety and security analysis of an AID system. This PoC was based on the literature on these systems [Deshpande et al., 2019, Bergenstal et al., 2016] and conducted by the authors as a first step toward empirical evaluation.

Agile Project Planning with the *SafeSecIoT Canvas*

To support the safety and security analysis process of an IoT system from its conception, the proposed agile project planning strategy places the completion of the *SafeSecIoT Canvas* as the initial RE activity. The canvas should be completed collaboratively by the stakeholders and technical team members involved in the project. In this PoC, the canvas was completed by the authors.

Figure 5.7 presents the *SafeSecIoT Canvas* for the AID system design. Building blocks 1–6 and 13–20 are general-purpose elements that address fundamental questions: i) why, ii) what, iii) who, and iv) when and how much. Blocks 6–9 and 10–12, on the other hand, are domain-specific elements focused on the "what" question, relating to the IoT system and security aspects. The project planning information captured using the canvas, as outlined in the proposed strategy, serves as input for the STPA-based analysis process, as will be detailed in Chapter 6.

Connecting *SafeSecIoT Canvas* to Defining the Purpose of the Analysis

Through the application of *SafeSecIoT Canvas*, as shown in Figure 5.7, the main system requirements were identified:

- R01 - The CGM shall measure blood glucose levels every $\leq$ 5 minutes.
- R02 - The algorithm shall calculate the insulin dose within 2 seconds after the reading.
- R03 - The pump shall administer the calculated dose with an accuracy of ±5%.
- R04 - The system shall suspend insulin when glucose $\leq$ 70 mg/dL.
- R05 - The system shall sound an alarm when the glucose level is < 70 mg/dL for more than 5 min.
- R06 - The system shall sound an alarm when the glucose level is > 250 mg/dL for more than 10 min.
- R07 - The system shall allow manual entry of blood glucose values in $\leq$ 30 s.
- R08 - The system shall record readings and doses for at least 30 days.
- R09 - The system shall remain operational for $\geq$ 99% of the time in 24 hours.
- R10 - The system shall transmit data between the sensor and app in $\leq$ 100 ms.
- R11 - The system shall encrypt communications with AES-256.
- R12 - The interface shall display current glucose within $\leq$ 2 clicks from the home screen.
- R13 - The system shall receive CGM data within $\leq$ 5 min via secure BLE.
- R14 - The system shall send dose commands to the pump in $\leq$ 200 ms.
- R15 - The pump shall store at least 300 units of insulin.
- R16 - The system shall operate for at least 24 hours without recharging.

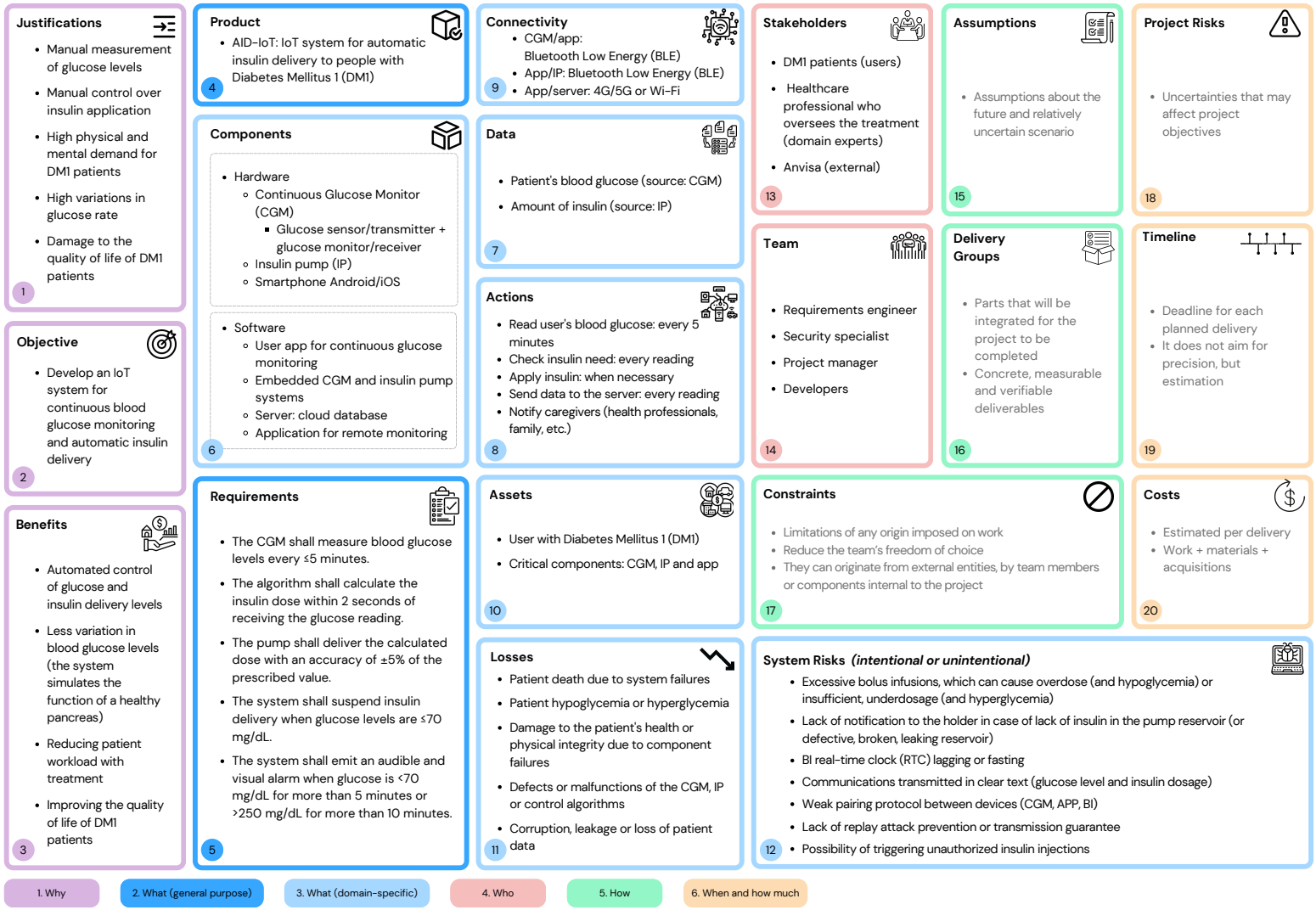


Figure 5.7: *SafeSecIoT Canvas* for agile project planning of an Automatic Insulin Delivery system.

By analyzing the requirements, together with the assets and other general and specific elements of the IoT domain present in the canvas, the *losses* (L) can be defined (in the planning stage with the *SafeSecIoT Canvas*) and refined and validated (in the context of an STPA-based analysis approach). The following *losses* were identified:

- L1: Loss of life or risk to the patient's life (acute episode of hypo/hyperglycemia, diabetic coma, etc.);
- L2: Loss or serious damage to AID system components;
- L3: Loss of mission (automated insulin delivery);
- L4: Loss or corruption of proportional or sensitive information;
- L5: Loss of connectivity between system components.

Linking Project Planning to Control Structure Modeling

An AID system comprises subsystems (or *components*) previously identified with the help of the *SafeSecIoT Canvas* (Figure 5.7: 6). Considering systems theory, a subsystem can be analyzed as another system (with more specific components and functions). Therefore, components can be successively expanded and detailed to the level necessary for each application, as exemplified in Figure 5.8.

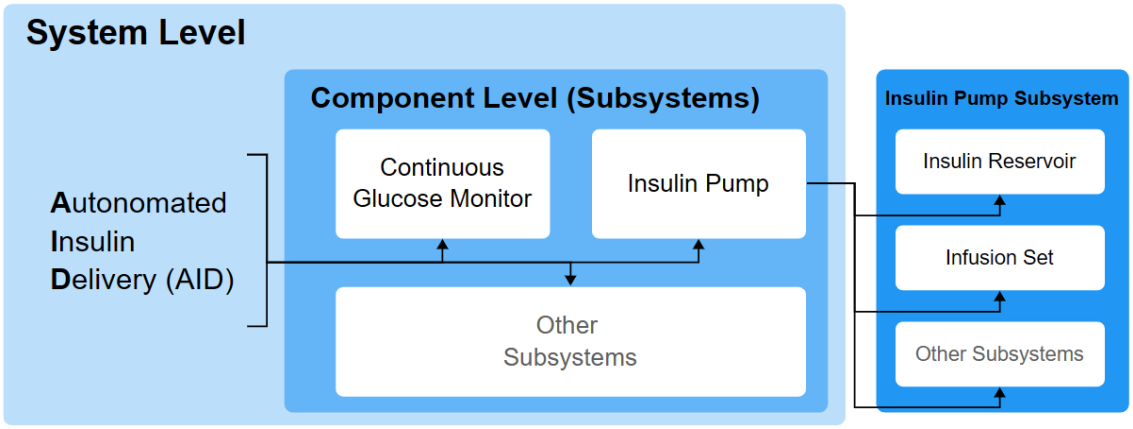


Figure 5.8: Automated Insulin Delivery subsystems.

Based on the specific domain elements mapped in the *SafeSecIoT Canvas*, the following components and associated responsibilities were identified:

- Continuous Glucose Monitoring (CGM):
  - R1: Measure blood glucose levels;
- Control Application (App):
  - R2: Check/calculate basal insulin requirements;
  - R3: Check/calculate insulin requirements for meal correction bolus;

- Insulin Pump (IP):
  - R4: Administer (apply) insulin.

Based on the responsibilities and components, the following control actions (CA) of the system were identified:

- CA1: Measure blood glucose;
- CA2: Release basal insulin;
- CA3: Release meal correction bolus;
- CA4: Deliver (administer) insulin.

In addition to supporting Stage 1, the IoT-specific components of the *SafeSecIoT Canvas* are directly used to model the AID system control structure. As illustrated in Figure 5.9, to link project planning (*SafeSecIoT Canvas*) with STPA-based safety and security analysis, the *components* are mapped as *controllers* or *controlled processes*, such as the CGM, the IP, and the smartphone (control application). Similarly, *actions* are mapped or derived as *control actions*, while *data* is used to derive *feedbacks*.

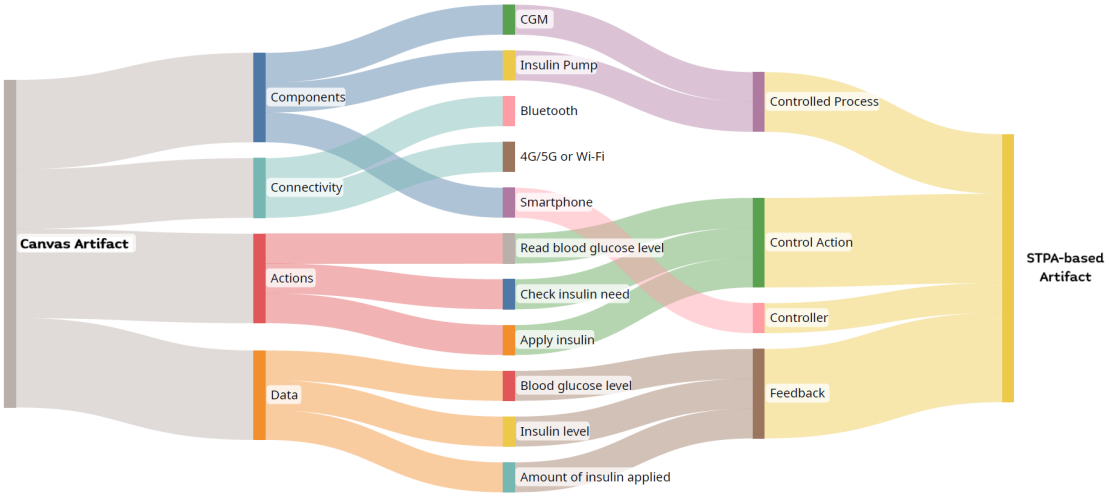


Figure 5.9: Traceability between artifact elements.

5.5 Chapter Summary

This chapter presented the *SafeSecIoT Canvas* model and demonstrated that the proposed strategy is based on mapping the building blocks of project planning (using the *SafeSecIoT Canvas*) to the subsequent elements of the critical IoT system safety and security analysis process (using a STPA-based approach, in this case, *STPA-SafeSecIoT*). To achieve this, the general-purpose elements provide the necessary inputs for defining the project's objectives, justifications, and essential requirements, while the domain-specific

elements (IoT, safety, and security) help further refine the system's scope. Thus, steps 1 and 2 of *STPA-SafeSecIoT* are directly supported and facilitated by the project planning carried out using the *SafeSecIoT Canvas*.

Based on what has been demonstrated about the proposed model, we seek to answer the following research question: "Can agile planning of critical IoT system projects be linked to STPA-based analysis and improve the quality of the safety and security RE process?"

1. The *SafeSecIoT Canvas* supports agile project planning with building blocks aligned with the objectives and needs of STPA-based safety and security analysis methods, such as *STPA-SafeSecIoT* [Veiga and Bulcão Neto, 2023];
2. This work formalizes a strategy for integrating the project planning artifact and the STPA-based safety and security analysis method, contributing to the analysis results;
3. The safety and security RE process and the artifacts developed are based on ISO/IEC/IEEE 15288. We seek to balance agile planning, based on a canvas artifact, with structured processes from the technical standard to improve the safety and security analysis of critical IoT systems;
4. The proposed model has the potential to improve communication between stakeholders and the technical team, including domain experts (in safety and security), raising the quality of planning information and, subsequently, the specified requirements.

The *SafeSecIoT Canvas* artifact and the proposed strategy address the gap in defining system scope and support the safety and security analysis of critical IoT systems. This methodology enables stakeholders and project teams to begin the analysis with prior system planning information, improving scope definition and facilitating the correct identification of losses, hazards, and threats. Consequently, safety and security requirements can be specified accurately, avoiding critical risks to system development and subsequent operation.

The presented PoC represents a preliminary validation step preceding empirical evaluation, providing initial results and insights. These results demonstrate the technical feasibility of the approach in supporting safety and security analysis through agile project planning. Lessons learned contributed to refining the *SafeSecIoT Canvas* artifact and improving its readiness for use in real projects.

CHAPTER 6

STPA-SafeSecIoT: An Extension of STPA for Safety and Security Analysis in Critical IoT Systems

This chapter presents an extension of the STPA method, proposed to support the alignment of safety and security in critical IoT systems: *STPA-SafeSecIoT*. The proposed method aims to enable: i) security analysis in parallel with safety analysis, allowing the identification of requirements related to these two perspectives, and also: ii) the specification of safety and security requirements for later use throughout the system development process. The objective of the proposed method is to address the alignment of safety and security from the system concept stage, supported by a canvas artifact, within an RE process for IoT systems, so that these requirements and related design decisions can be used in subsequent stages, such as design, construction, and testing.

6.1 Alignment of Safety and Security Requirements

Safety and security are essential requirements in critical IoT systems and must be addressed from the early stages of the development life cycle [Kavallieratos et al., 2020a]. Safety is the ability of the system to function without failures that could harm people or the environment, addressing accidental risks without malicious intent; while security is the ability of the system to defend itself against accidental or deliberate intrusions, focusing on risks of malicious intent [Lisova et al., 2019b, Lyu et al., 2019]. A system is considered critical because any failure can result in serious or fatal consequences. In addition, safety and security concerns are essential, including protection against cyber attacks and the integrity and confidentiality of the data generated and manipulated.

Systems have become increasingly complex and networked. This growing complexity requires new methods for systems engineering. In this regard, [Leveson and Thomas, 2018] developed *System Theoretic Process Analysis* (STPA), which is a well-established proactive risk analysis method that assesses system safety

issues by identifying potential causes of accidents, allowing risks to be eliminated or controlled. However, STPA does not address the alignment of safety and security requirements, which is necessary for critical IoT systems.

The next section presents this research proposal for aligning safety and security requirements for critical IoT systems. The proposal in question uses Systems Engineering methodologies, such as systems analysis and concepts related to the abstraction of complex systems, relating these elements to the analysis and specification activities of the RE process, in the form of a method that supports the definition of its safety and security requirements, together with the relationships between these requirements.

6.2 *STPA-SafeSecIoT* method: steps and tasks

Safety and security are essential requirements for avoiding losses in complex and critical software-controlled systems. In this sense, this work proposes an extension of the STPA method [Leveson and Thomas, 2018] (presented in Chapter 2), which aims to align safety and security requirements for critical IoT systems. STPA is considered a formal method for safety analysis which, within the scope of this proposal, is extended to perform joint safety and security analysis based on the identification of hazards/threats.

The objective of this method is the joint analysis of safety and security to specify requirements that can prevent not only system losses caused by hazards and/or unintentional accidents, but also system losses introduced by unknown and/or intentional threats, such as malicious individuals or organizations.

6.2.1 Step 1: Defining the Purpose of the Analysis

The first step of the proposal presented in this paper extends Step 1 of the STPA (to define the purpose of the analysis), based on the *SafeSecIoT Canvas* to enable joint analysis (or co-analysis) of system safety and security. It has four main tasks:

- Define the system and system boundaries (project planning and scope);
- Identify unacceptable losses;
- Identify hazards and threats at the system level;
- Identify safety and security constraints at the system level.

Figure 6.1 presents an overview of Step 1, showing the actions performed and the outputs produced in this step, artifacts for safety and security analysis.

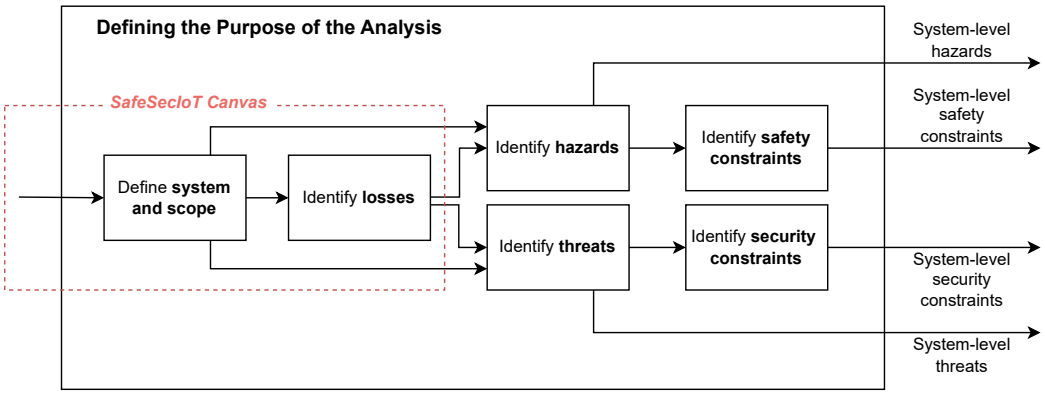


Figure 6.1: Step 1 - Defining the purpose of the analysis.

Task 1.1. Define the System and its Scope

Before beginning the safety and security analysis process, it is first necessary to define the system to be analyzed and the limits of this system, i.e., its scope. This activity of the proposed approach aims to fill a gap in the STPA, which is the identification of the subsystems/components of the analyzed system and a high-level description of its scope, using the *SafeSecIoT Canvas* for this purpose.

Both in the STPA method and in approaches based on it, there is a lack of information on the scope of the system that can hinder or limit the understanding of hazards and threats, as well as the modeling of the system's control structure. To perform Step 2, for example, it is necessary to identify the subsystems to be analyzed, down to the required level of granularity. The identification of these components of an IoT system, together with the essential requirements that make up the scope, not only allows for a better understanding of the system as a whole, but also provides a basis for modeling the control structure.

In systems engineering, one way to define the boundaries of the system for analysis purposes is to include in this scope the parts of the system over which designers have some control [Leveson and Thomas, 2018]. In this sense, in order to guide, facilitate, and support the process of defining the system scope, as well as eliciting the essential requirements of the system, we propose to extend this task using a metamodeling-based approach and an artifact for planning critical IoT system projects, the *SafeSecIoT Canvas*, as presented in Chapters 4 and 5.

The *SafeSecIoT Canvas* provides structured input to support the definition of losses, hazards, threats, and system-level safety and security constraints, acting as an initial step in the RE process. By capturing strategic information about the project context, IoT-specific characteristics, and preliminary safety and security concerns, the canvas facilitates the identification of critical elements that may lead to unacceptable

outcomes. This structured representation reduces ambiguity and ensures that relevant aspects are explicitly considered from the early stages of system design. In this sense, the canvas complements and integrates with STPA-based approaches, providing a systematic foundation for the subsequent derivation of safety and security requirements, as well as for the formalization of preventive and protective constraints.

Task 1.2. Identify Losses

To suit the context of this work, we extend the definition of *loss* proposed by [Leveson and Thomas, 2018]: *"A loss involves something of value to stakeholders or the system, which may include loss of life or injury, property damage, environmental damage, mission loss, loss of reputation, loss or leakage of confidential data/information, or any other loss related to safety and security vulnerabilities that is unacceptable to stakeholders."*

This definition is important, since different terms are used, depending on the domain, to identify the objective of a safety/security analysis. Examples of this are the terms "accident", "adverse event", "incident", "attack", among others. To avoid this problem, we have adopted the term "loss" in this work. Therefore, the objective of the proposed approach is to arrive at the requirements necessary to avoid safety and security-related losses, identified in the initial stage of the analysis.

The main objective of the STPA method, on which this part of the proposal is based, is to avoid losses, so it can be used to identify any type of loss that is unacceptable to stakeholders [Leveson and Thomas, 2018]. It should also be noted that all artifacts produced from this point in the analysis will be defined to maintain traceability to the losses, so that the results produced in the following stages can be traced back to one or more identified losses. This factor aims to contribute to activities such as prioritizing requirements and even design decisions, since it should be possible to identify and rank which losses will have the most impact on the system.

A general approach to identifying losses in systems involves three essential steps [Leveson and Thomas, 2018], which are extended in this work:

- Identify the stakeholders or assets of the system;
- Identify the values and objectives that stakeholders or assets expect from the system;
- Translate each of these values into an unacceptable loss from the point of view of a stakeholder or asset.

Examples of unacceptable losses to stakeholders may include:

- Loss of life or injury to persons;

- Loss of or damage to the system;
- Loss or damage to the environment or objects external to a system;
- Loss of confidential information;
- Loss of control over the system.

A loss should not refer to individual components or specific causes. On the other hand, losses are broad enough to involve aspects of the environment that are not directly controlled by a system designer. In addition, any consideration or assumption made by an analyst during the identification of losses must be documented, for example, the reason why a particular loss was explicitly excluded from the analysis.

Task 1.3. Identify System-level Hazards and Threats

At this stage, as an essential part of the STPA extension, the proposed approach aims to address safety and security concerns in an integrated manner, placing equal importance on both factors in the system design process and subsequent definition of requirements. In this sense, events that may cause safety-related losses are classified as *hazards*, while those that may cause security-related losses are defined as *threats* [Lautieri et al., 2005].

In general, hazards are considered to be those that arise from unintentional errors or mistakes, while threats are those that come from deliberate actions intended to cause some kind of impact [Gromule et al., 2017]. The study by [Yang and Qu, 2016] further argues that hazards are generally caused by internal and external factors, while threats mainly result from external sources, although there may be some internal assistance.

One of the contributions of the approach proposed in this work is the joint analysis of safety and security from the system design stage and the tracking of identified requirements to the possible hazards and threats they are intended to mitigate. For this reason, immediately after identifying possible losses, an analysis is performed to identify hazards and threats at the system level, initiating the co-analysis of system safety and security.

Therefore, this step of the analysis prescribes the identification of hazards and threats at the system level, and the association of hazards and threats with the losses identified in the previous step. In general, a hazard or threat can lead to one or more losses, and each hazard/threat must be traceable to the resulting losses.

Some criteria, based on STPA, can be used to guide the activity of defining hazards/threats at the system level:

- Hazards/threats are states or conditions of the system, i.e., they should not describe detailed causes at the component level (such as brake failure, insufficient battery, etc.);

- A hazard/threat will lead to a loss in a worst-case scenario, i.e., a hazard/threat in isolation may not lead to a loss;
- Hazards/threats should describe states or conditions to be prevented, i.e., states that the system should never enter.

A common mistake in identifying hazards/threats is to confuse them with their causes [Leveson and Thomas, 2018]. Hazards/threats should refer to the system as a whole, not to a specific component of the system. Another factor that indicates problems in defining hazards/threats is a large number and/or excessive detail of hazards/threats. If necessary, broader hazards/threats can be refined later into more detailed descriptions.

Another problem is the use of ambiguous terms or recursion in the definition of hazards/threats [Leveson and Thomas, 2018]. There is often a natural tendency to use the term "unsafe" in the definition of a hazard/threat. However, the use of such terms makes the definition of the hazard or threat very vague, not allowing the actual unsafe condition that should be referred to be specified. Therefore, the conditions or states of the system that make it unsafe must be clearly indicated.

In some cases, hazards/threats may be confused with failures, especially at the component level. Hazards/threats are defined at a high level of abstraction, and it is not possible to differentiate their causes from technical failures, design errors, problems with requirements, among others. This detailing of causes will be carried out later in the safety and security analysis process (Step 3). In order to avoid these problems in conducting the analysis, this proposal also analyzes terminology and techniques for analyzing hazards and threats, which can serve as a reference and theoretical support to assist in carrying out this activity.

Task 1.4. Identify System-level Safety and Security Constraints

A system-level constraint specifies the conditions or behaviors that must be enforced to prevent the occurrence of previously identified hazards (safety-related) and threats (security-related), thereby ultimately avoiding losses. These constraints are systematically derived from hazards and threats, representing their formal negation in the context of system design and operation.

Traceability between constraints and hazards/threats may not be one-to-one, as is the case for losses, since the same hazard/threat may be related to several losses. Therefore, for each hazard/threat there may be several constraints (if the hazard/threat has been refined). This means that there is a many-to-many relationship between hazards/threats, both with losses and with constraints.

As mentioned earlier, hazards/threats can be refined, if necessary, to help identify more specific restrictions, but still at the system level. This should be done when there is

a subsystem/component that needs to be controlled in order to avoid this hazard/threat at the system level.

In summary, Step 1 involves defining the scope of the system and requirements using the *SafeSecIoT Canvas* and identifying important information for safety and security analysis: first, the losses, then the hazards/threats that can lead to these losses, and finally, the system-level constraints that specify the conditions or behaviors necessary to avoid the corresponding hazards/threats and associated losses.

6.2.2 Step 2: Control Structure Modeling

The second stage of the proposed approach follows the general process of modeling the control structure established in STPA. However, its main advancement lies in extending the analysis beyond hazards to also encompass threats, thereby broadening the scope to jointly address both safety and security concerns. This integrated perspective overcomes one of the main limitations of traditional STPA, which is typically focused on safety alone, and strengthens the capability to capture interdependencies between safety and security aspects in critical IoT systems.

A general overview of the control structure and the essential concepts required for understanding this stage are presented in Subsection 2.3.2. Subsequently, the necessary steps for modeling the control structure are described, along with the information and artifacts generated from this stage, illustrated through representative examples.

Task 2.1. Identify Components of the Control Structure

Control structures use abstraction to manage system complexity. This means that, instead of listing each component or subsystem individually, the analysis can start at a more abstract level, assuming that the controller (or controlled process) can be a set of components/subsystems, thus generating fewer hierarchical levels in the control structure. In this way, the task of modeling the control structure can begin at a high level, according to the needs of the system, and the necessary details are added interactively.

One possibility is to start the control structure with the basic subsystems necessary to impose the constraints and avoid the hazards/threats identified earlier. In this way, the constraints identified in the previous step will guide the definition of which subsystems are necessary for modeling the control structure. Figure 6.2 presents a high-level control structure template.

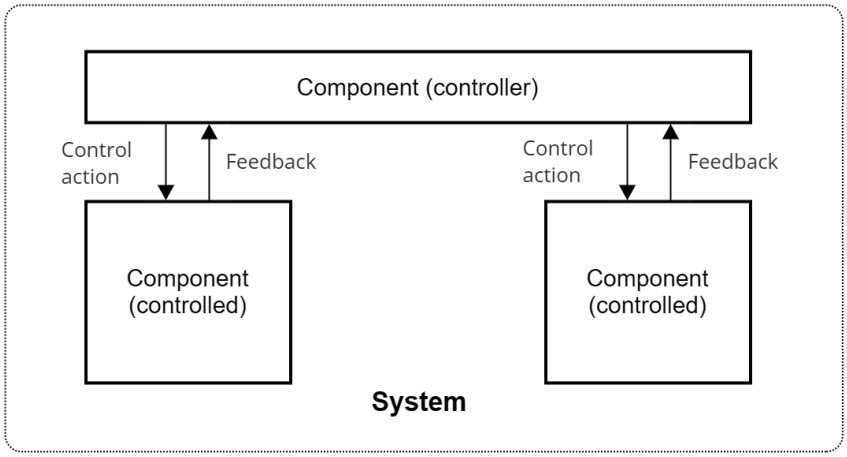


Figure 6.2: Step 2 - High-level control structure template.

Task 2.2. Define Component Responsibilities

Once the controllers have been identified (and refined, if necessary), responsibilities can be defined for each of these controllers by refining the system-level safety/security constraints identified in Step 1. One question that can guide this definition is: what does each entity need to do so that collectively the high-level system constraints are guaranteed?

Defining responsibilities provides an overview of which constraints are affected together, allowing an initial view of which hazards and threats are directly related. This information can be used to help identify the relationships between safety and security requirements.

Task 2.3. Derive Control Actions

Just as responsibilities were derived from safety and security restrictions at the system level, control actions can be derived from responsibilities. Thus, control actions for each controller can be defined based on the identified responsibilities.

Task 2.4. Derive Feedbacks

Feedback can be derived from control actions and responsibilities, identifying the process models that controllers need to make decisions.

6.2.3 Step 3: Identification of Unsafe/Unsecured Control Actions and Safety and Security Requirements

An overview of Step 3 is presented in Figure 6.3.

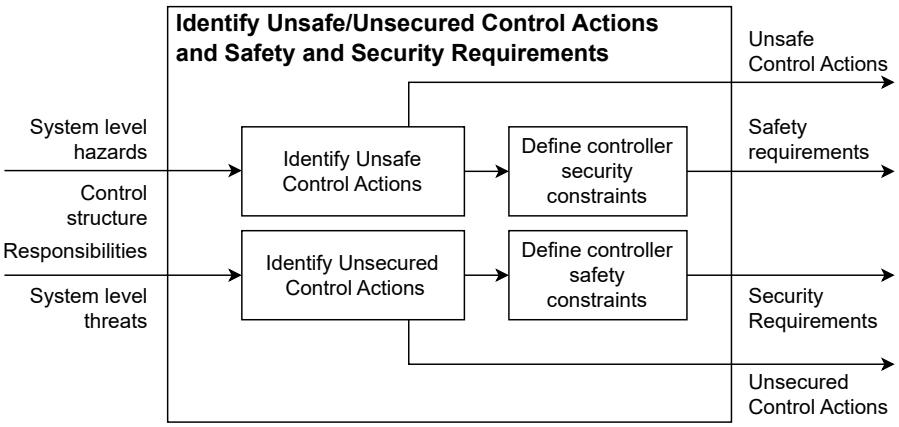


Figure 6.3: Step 3 - Identify UCAs and Safety and Security Requirements.

Task 3.1. Identify Unsafe/Unsecured Control Actions

An Unsafe/Unsecured Control Action (UCA) is a control action that, in a particular context and in a worst-case scenario, will lead to a hazard or threat. In the context of this proposal, we classify UCAs into two categories: unsafe when they refer to hazards (safety) or unsecured when they refer to threats (security).

UCAs are defined for each controller and its control actions. In STPA, there are four ways in which a control action can be considered unsafe, which have been adapted in this proposal to the context of safety and security, into four types:

1. Failure to provide the control action, leading to a hazard/threat.
2. Providing the control action, leading to a hazard/threat.
3. Providing a potentially safe/secure control action, but too early, too late, or in the wrong order.
4. The control action lasts too long or is interrupted too early (for continuous, non-discrete control actions).

UCAs must be associated (in their description) with a hazard or threat (identified previously) in order to maintain the traceability of the analysis. There may be more than one unsafe control action for each of the four types mentioned above, referring to each control action and hazard/threat analyzed.

A UCA must specify, in its description, the context in which the control action is unsafe, and this context is critical. The context allows us to understand why a control action is unsafe: if a control action were always unsafe, then engineers would never include it in the system design. Any relevant context can be referenced in a UCA, including environmental conditions, states of the controlled process, states of the controller, previous actions, and parameters. Generally, the context is identified after the terms “when,” “while,” “during,” among others.

The description of a UCA can be divided into five parts:

- Source: the element from which the control originates;
- Type: defines the type of UCA (presented above);
- Control Action: the control action itself, derived from the control structure, which is being analyzed as unsafe;
- Context: must specify the state or condition in which the UCA occurs;
- Link to the hazard/threat: identification of the hazard or threat to which the UCA is linked.

Thus, we have the following specification model for a UCA:
UCA: <Source> <Type> <Control Action> <Context> <Association with danger or threat at the system level>

Task 3.2. Define Component-level Safety and Security Requirements

Once the UCAs have been identified, it is then possible to define the safety and security requirements associated with the controllers. Controller requirements are different from the constraints identified in Step 1, which are system-level safety and security constraints. A controller requirement specifies the controller behaviors that are necessary to prevent UCAs.

From the identified unsafe or unsecured control actions, the safety and security requirements of the critical IoT system are derived. These requirements define the behaviors that must be enforced to prevent the occurrence of such unsafe or unsecured actions. The safety and security requirements specification model:
Safety/Security Requirement = <Source> & < Behaviours that need to be satisfied> & Control Action> & <Constraint/Context>

Thus, the inputs for Stage 3 of the process are the hazards/threats at the system level (as well as their refinements), the control structure, and the responsibilities. And as output from this stage, we have the UCAs and the safety and security requirements of the system.

6.2.4 Step 4: Identification of Loss Scenarios

An overview of Step 4 is presented in Figure 6.4.

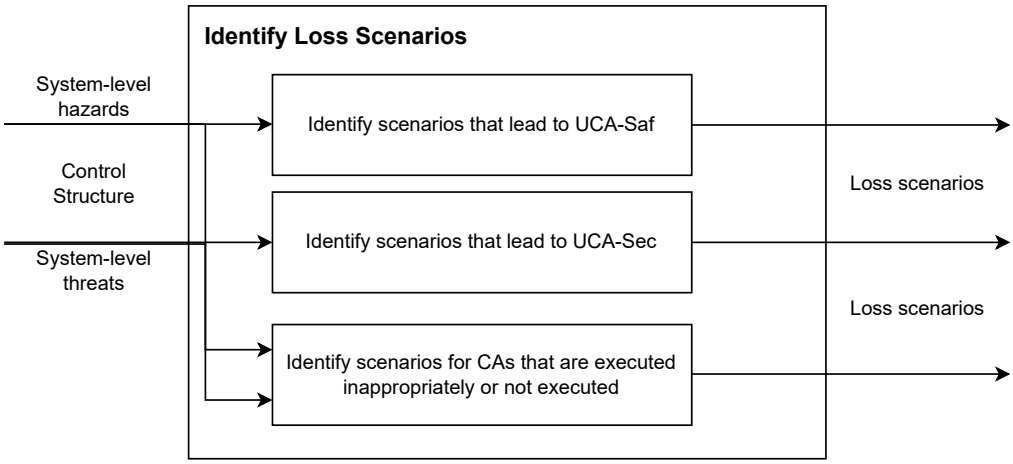


Figure 6.4: Step 4 - Identify Loss Scenarios.

Adapting the definition from [Leveson and Thomas, 2018], a loss scenario describes the causal factors that can lead to UCAs and hazards/threats. There are two types of loss scenarios associated with control actions that can be considered (illustrated in Figure 2.9 in Chapter 2):

- Why do UCAs occur?
- Why are UCAs performed inadequately or not performed at all, leading to hazards/threats?

The first type of loss scenario that can occur is related to UCAs. There are two causes that lead to UCAs: i) unsafe controller behavior and ii) inadequate feedback and other inputs. In this context, scenarios can be identified in which a controller will issue UCAs or feedback that leads to these UCAs.

There are four general reasons why a controller may provide (or fail to provide) an unsafe control action:

1. Failures involving the controller (mainly for physical controllers);
2. Inappropriate control algorithm;
3. Unsafe control input (from another controller);
4. Inadequate process model:
    - The controller receives incorrect feedback or information;
    - The controller receives correct feedback or information but interprets it incorrectly or ignores it;
    - The controller does not receive feedback or information when necessary;
    - Feedback or information necessary for the controller does not exist.

The description of scenarios for the occurrence of UCAs, in addition to describing the occurrence that leads to a UCA, identifies which UCA it refers to as well as which hazard/threat it is related to.

In a scenario that includes security-related causes, an additional possibility must be considered: identifying how the specified feedback or other information could be affected by a malicious agent (who is carrying out some type of attack). More specifically, it is necessary to identify how incorrect feedback or information could be injected, falsified, intercepted, or disclosed by an adversary, for example, based on a threat model such as STRIDE [Shostack, 2014].

A second type of scenario that can occur, and lead to dangers/threats and consequently to losses, is not linked to the existence of UCAs, but rather to control actions that are performed inadequately or are not performed at all. Some factors can be taken into account when creating scenarios that indicate the reasons why control actions are performed inadequately or not performed at all: i) factors that affect the control path; and ii) factors that affect the controlled process.

To identify loss scenarios, system-level hazards/threats (as well as their refinements), the control structure, and UCAs are used as inputs to the process. The output of this process is the identification of scenarios that lead to the implementation of UCAs and loss scenarios caused by the non-execution or incorrect execution of control actions. To identify loss scenarios, system-level hazards/threats (as well as their refinements), the control structure, and UCAs are used as inputs to the process. The output of this process is the identification of scenarios that lead to the realization of UCAs and loss scenarios caused by the non-execution or incorrect execution of control actions.

6.2.5 Traceability between information produced in the process

Figure 6.5 shows the traceability that must be maintained between the information (outputs) of each stage of the safety and security analysis process presented.

As shown in the figure, in Step 1, hazards/threats at the system level are defined based on losses and are directly associated with them. Based on the hazards/threats, safety and security constraints at the system level (SC-Saf and SC-Sec) are defined. In Step 2, components are defined based on components/subsystems and have control actions and responsibilities. Responsibilities are derived from system-level constraints, and from these, control actions and feedback are derived. In Step 3, the UCA-Saf and UCA-Sec are defined, which are directly related to the hazards/threats defined in Step 1, and from these, the controller restrictions are derived. Finally, in Step 4, scenarios based or not based on UCAs are defined, associated respectively with the UCAs and the hazards/threats.

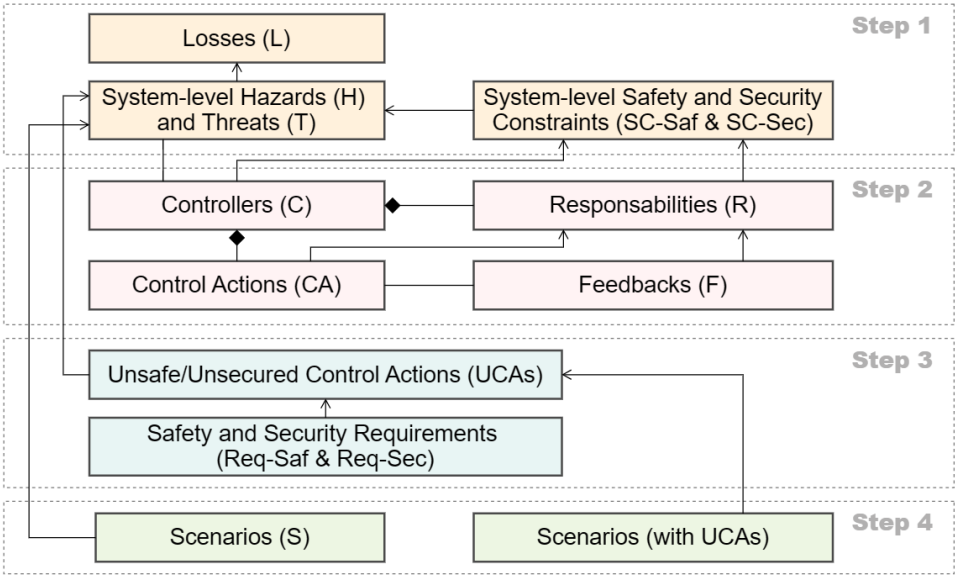


Figure 6.5: Traceability between the information produced in the safety and security analysis process.

6.3 Proof of Concept: analysis and specification of safety and security of an AID system

Step 1: Defining the Purpose of the Analysis

Task 1.1. Define the System and its Scope

As presented in the PoC in Chapter 5, Section 5.4, the *SafeSecIoT Canvas* is used for project planning and defining the scope and essential requirements of the IoT system. Figure 5.7 shows the canvas filled with the information that will be used as inputs for the STPA-based analysis process.

Task 1.2. Identify Losses

According to the planning information and requirements identified in the *SafeSecIoT Canvas* (Figura 5.7), the following losses for the system were validated:

- L-1: Patient experiences severe hypoglycemia or hyperglycemia (or death).
- L-2: System fails to maintain therapeutic efficacy over time.
- L-3: Loss of confidence in the AID system by patients or caregivers.
- L-4: Unauthorized access or manipulation of system functions or data.
- L-5: Violation of patient privacy or data protection regulations.

These losses cover the full spectrum of possible consequences (clinical, functional, social, legal, and cybersecurity) and provide the basis for deriving hazards, threats, and safety and security constraints according to the STPA-based approach.

Task 1.3. Identify System-level Hazards and Threats

Next, in the safety and security analysis process, *hazards* (H) and *threats* (T) are identified at the system level. The hazards and threats identified are based on the losses themselves and also on the risks identified in the *SafeSecIoT Canvas* artifact. Hazards are directly linked to functional or clinical failures, while threats reflect cybersecurity and physical security scenarios for this type of critical system. In both cases, these are situations that, in the worst case, can lead to one or more losses. The following hazards and threats were identified:

- H-1: The AID system administers an incorrect dose of insulin due to inaccurate or invalid CGM readings [L-1][L-2]
- H-2: The AID system makes dosing decisions based on outdated data due to loss of connectivity between CGM, controller, and pump [L-1][L-2]
- H-3: The AID system fails to detect an occlusion or failure in the infusion set in a timely manner, resulting in interruption or failure of insulin delivery [L-1][L-2]
- H-4: The AID system malfunctions during high-risk user activities (e.g., intense exercise, driving, sleeping) [L-1][L-2][L-3]
- H-5: The AID system does not adapt the control algorithm to rapid physiological changes in the user (e.g., exercise, stress, eating) [L-1][L-2]
- T-1: The AID system allows unauthorized remote access to the insulin pump, enabling malicious commands [L-1][L-2][L-3][L-4]
- T-2: Data transmitted between CGM, controller, and pump is intercepted, altered, or reused by attackers (man-in-the-middle or replay attack) [L-1][L-2][L-4]
- T-3: The AID system contains software/firmware vulnerabilities that can be exploited to alter its operation or disable security mechanisms [L-1][L-2][L-3][L-4]
- T-4: The AID system does not implement sufficiently strong authentication or access control, allowing misuse or privilege escalation [L-1][L-2][L-3][L-4]
- T-5: The user's personal and clinical data is accessed or misused by unauthorized third parties [L-3][L-4][L-5]

Task 1.4. Identify System-level Safety and Security Constraints

Based on the identified hazards and threats, safety and security restrictions are derived at the system level, aiming to mitigate the realization of risks and, in the worst case, possible losses:

- SC-Saf-1: The AID system should administer insulin doses only based on validated blood glucose values, rejecting readings outside the acceptable physiological range (40–400 mg/dL) [H-1]
- SC-Saf-2: The AID system should suspend automatic insulin delivery and alert the user if CGM data is not updated within 5 minutes [H-2].
- SC-Saf-3: The AID system should detect infusion failures within 15 minutes and immediately notify the user [H-3].
- SC-Saf-4: The AID system should have operating modes adapted to critical contexts (sleep, exercise, driving), adjusting safety limits as configured [H-4].
- SC-Saf-5: The control algorithm must adjust the insulin rate in response to rapid glycemic variations ($\geq$ 2 mg/dL/min) in less than 5 minutes [H-5].
- SC-Sec-1: The AID system must restrict remote control commands to authenticated and encrypted communication with unique keys per device [T-1].
- SC-Sec-2: All communications between CGM, controller, and pump must use end-to-end encryption and replay protection [T-2].
- SC-Sec-3: The AID system must apply continuous integrity verification and support secure, digitally signed updates [T-3].
- SC-Sec-4: The AID system must implement multi-factor authentication and separation of privileges at all access levels [T-4].
- SC-Sec-5: The AID system must store and transmit personal data only in encrypted form, with the explicit consent of the user [T-5].

System-level constraints identify high-level countermeasures that the system must implement to avoid the respective hazards and threats. Based on the control structure modeling, a more detailed analysis of the system's control components is performed, identifying control actions and possible unsafe control actions and, subsequently, the respective controller-level constraints.

Step 2: Control Structure Modeling

Task 2.1. Identify Components of the Control Structure

Based on the components identified in the *SafeSecIoT Canvas* (Figura 5.7, item 11), the following subsystems were identified to be included in the control structure:

- C-1: Control application (app);
- C-2: Continuous glucose monitor (CGM);
- C-3: Insulin pump (IP);
- C-4: Server;
- C-5: Patient;

- C-6: Caregiver or health professional.

Task 2.2. Define Component Responsibilities

Table 6.1 presents the responsibilities identified for each component/controller of the system.

Task 2.3. Derive Control Actions

The system control actions can be derived and refined from the actions identified in the *SafeSecIoT Canvas* (Figure 5.7, item 8) together with the analysis of the responsibilities of the system components identified in the previous task. In this case, we identified the following as the main control actions:

- Read patient's blood glucose: (C2 $\rightarrow$ C5);
- Release insulin bolus: (C1 $\rightarrow$ C3);
- Deliver insulin: (C3 $\rightarrow$ C5);
- Send data to the server (C1 $\rightarrow$ C4);
- Notify caregivers or health professionals: (C4 $\rightarrow$ C6).

Table 6.1: Responsibilities of control structure components in the AID system

Component (ID)	**Main Responsibilities (with associated constraints)**
C-1: Control Application (App)	- Execute the control algorithm to calculate insulin doses [SC-Saf-1]. - Validate and process received data (CGM, manual inputs) [SC-Saf-1]. - Ensure secure and reliable communication with CGM, insulin pump, and server [SC-Sec-2][SC-Sec-3]. - Display information, alerts, and recommendations to patient/caregiver [SC-Saf-4]. - Implement fallback (manual mode) in case of failures [SC-Saf-4][SC-Saf-5].
C-2: Continuous Glucose Monitor (CGM)	- Continuously monitor blood glucose levels [SC-Saf-1]. - Transmit real-time readings to the app [SC-Sec-2]. - Ensure accuracy and integrity of data [SC-Saf-1][SC-Sec-2]. - Detect and report calibration errors, reading failures, or signal loss [SC-Saf-1][SC-Saf-5].
C-3: Insulin Pump (IP)	- Deliver insulin according to app instructions [SC-Saf-1]. - Monitor status (battery, occlusion, infusion failure) [SC-Saf-3]. - Report operational state to the app [SC-Sec-2][SC-Saf-3]. - Execute only authorized commands [SC-Sec-1][SC-Sec-4]. - Activate fail-safe (stop infusion in unsafe conditions) [SC-Saf-3][SC-Saf-4].
C-4: Server	- Securely store clinical data and history [SC-Sec-5]. - Support remote monitoring and parameter adjustments [SC-Sec-2][SC-Sec-4].
C-5: Patient	- Correctly use the system (sensor, pump, basic maintenance) [SC-Saf-5]. - Provide accurate manual inputs (e.g., carbohydrates, blood glucose) [SC-Saf-1]. - Respond to alarms and notifications [SC-Saf-4][SC-Saf-5]. - Report failures or unexpected behavior [SC-Saf-5].
C-6: Caregiver or Health Professional	- Supervise system use in vulnerable patients [SC-Saf-5]. - Analyze historical data and adjust clinical parameters [SC-Saf-1][SC-Saf-2]. - Intervene in critical situations or when the patient does not respond to alarms [SC-Saf-4][SC-Saf-5].

Task 2.4. Derive Feedbacks

Table 6.2: Main feedbacks between components in the AID system

Feedback Source → Target	**Feedback Information**
C2: CGM → C1: Control App	- Continuous glucose values. - Signal quality and calibration status. - Reading errors or sensor disconnection.
C3: Insulin Pump → C1: Control App	- Pump status (battery, reservoir level). - Infusion status and history of delivered doses. - Alarms (occlusion, delivery failure).
C1: Control App → C5: Patient	- Confirmation of insulin delivery. - Visual and/or audible alarms (low battery, occlusion, malfunction). - Insulin dose recommendations and bolus suggestions. - Alerts on hypo/hyperglycemia risks. - Warnings about communication failures or need for manual action.
C4: Server → C1: Control App	- Confirmation of data synchronization. - Historical data access. - Updates of clinical parameters.
C4: Server → C6: Caregiver or Health Professional	- Patient historical data (glucose trends, insulin delivery). - Alerts about anomalies or critical events.
C5: Patient → C1: Control App	- Manual input of carbohydrate intake, capillary glucose, physical activity, or symptoms.
C6: Caregiver / Health Professional → C4: Server	- Adjustments of clinical parameters (basal rate, glucose limits, alarm thresholds). - Clinical supervision reports.

Define the Control Structure

Based on the analysis performed, Figure 6.6 shows the control structure (at a high level) defined for the AID system.

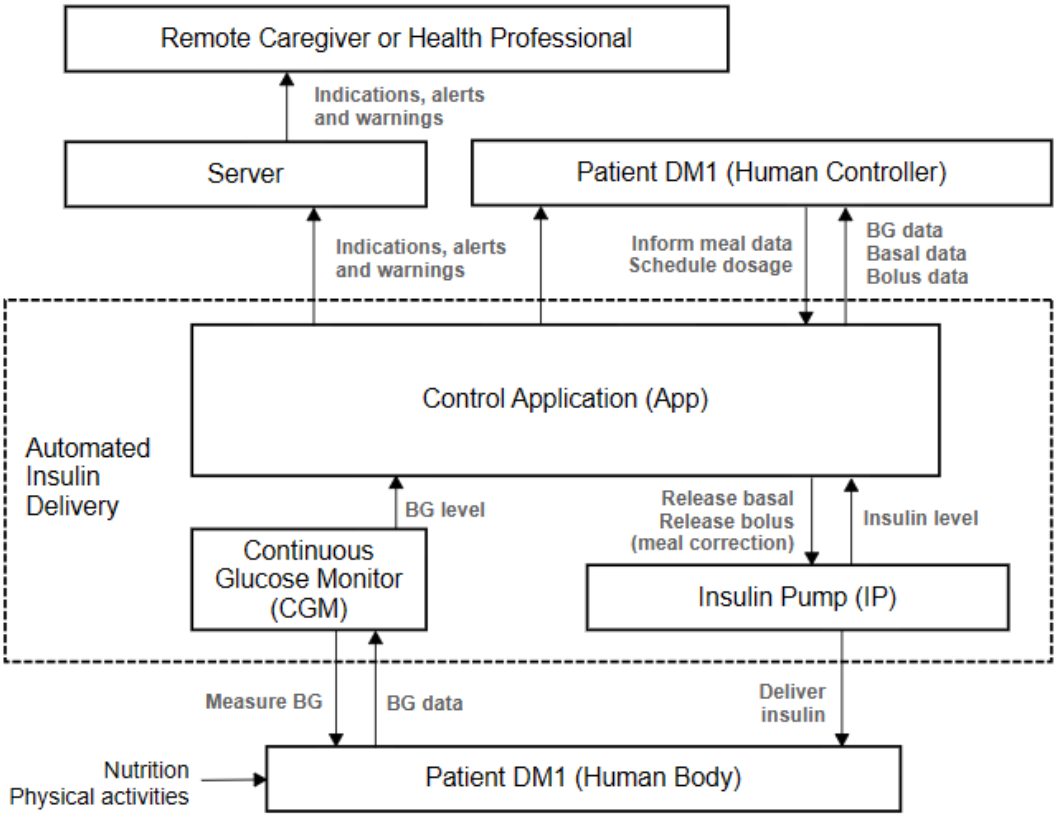


Figure 6.6: Control structure for the AID system.

Step 3: Identification of Unsafe/Unsecured Control Actions and Safety and Security Requirements

Task 3.1. Identify Unsafe/Unsecured Control Actions

Unsafe Control Actions (UCA-Saf):

- UCA-Saf-1: CGM does not provide blood glucose reading when Control App requests it during post-meal monitoring (ideally 2 hours after meal) [H-1, H-2]
- UCA-Saf-2: CGM provides inaccurate or invalid blood glucose reading during rapid glucose change (e.g., post-exercise or post-meal) [H-1]
- UCA-Saf-3: CGM reading is delayed by more than 5 minutes after measurement time requested by Control App [H-2]
- UCA-Saf-4: Control App does not send insulin bolus command when glucose exceeds 180 mg/dL after meal [H-1, H-5]
- UCA-Saf-5: Control App sends insulin bolus command when glucose is below 140 mg/dL or falling (e.g., during exercise) [H-1]
- UCA-Saf-6: Control App sends insulin bolus command more than 30 minutes before expected post-meal peak [H-5]

- UCA-Saf-7: Control App sends insulin bolus command more than 30 minutes after post-meal peak, increasing hyperglycemia risk [H-5]
- UCA-Saf-8: Insulin Pump does not deliver insulin after receiving command from Control App within 5 minutes [H-1, H-3]
- UCA-Saf-9: Insulin Pump delivers insulin without a valid command (manual override or software error) [H-1]
- UCA-Saf-10: Insulin Pump delivers insulin more than 5 minutes after bolus command, missing intended glucose control window [H-1, H-3]
- UCA-Saf-11: Insulin Pump stops basal insulin infusion before completing programmed dose [H-1, H-3]
- UCA-Saf-12: Insulin Pump continues basal infusion beyond programmed duration, causing overdose risk [H-1, H-3]
- UCA-Saf-13: Control App does not send patient data to Server during connectivity failure, preventing remote monitoring [H-2, H-4]
- UCA-Saf-14: Control App sends corrupted or inconsistent data to Server due to sensor error [H-2, H-5]
- UCA-Saf-15: Server does not notify caregiver about critical hypoglycemia or hyperglycemia event [H-4]
- UCA-Saf-16: Server sends false alarms or reassurance messages, causing caregiver confusion [H-4]
- UCA-Saf-17: Server delays notification to caregiver beyond safe intervention window [H-4]

Unsecured Control Actions (UCA-Sec):

- UCA-Sec-1: Insulin Pump delivers insulin based on forged commands from attacker instead of Control App [T-1]
- UCA-Sec-2: Control App sends insulin bolus command after being hijacked by attacker [T-1, T-3]
- UCA-Sec-3: CGM provides manipulated glucose readings due to man-in-the-middle attack [T-2]
- UCA-Sec-4: Control App sends delayed insulin bolus command due to replay attack [T-2]
- UCA-Sec-5: Control App sends repeated insulin bolus command due to replay attack [T-2]
- UCA-Sec-6: Server sends false alarms or reassurance messages to caregivers due to attacker compromise [T-2, T-3]
- UCA-Sec-7: Control App does not enforce authentication for Pump access, allowing unauthorized takeover [T-1, T-4]

- UCA-Sec-8: Server does not enforce access control to patient records, allowing unauthorized disclosure [T-5]
- UCA-Sec-9: Control App transmits patient data without encryption, allowing interception of sensitive information [T-2, T-5]
- UCA-Sec-10: Server provides delayed or manipulated feedback to caregivers due to compromised infrastructure [T-2, T-3]

Task 3.2. Define Component-level Safety and Security Requirements

Below, we show some examples of requirements based on the UCAs defined above:

- Req-Saf-1: CGM must provide accurate blood glucose readings within $\pm$10 mg/dL whenever requested by Control App, including during post-meal monitoring (up to 2 hours after meal) [UCA-Saf-1]
- Req-Saf-2: CGM must detect rapid glucose changes (e.g., post-exercise, post-meal) and provide valid readings within 5 minutes [UCA-Saf-2]
- Req-Saf-3: CGM readings must be transmitted to Control App within 5 minutes of measurement [UCA-Saf-3]
- Req-Saf-4: Control App must send insulin bolus command when glucose exceeds 180 mg/dL after meal, considering meal detection logic and patient-specific parameters [UCA-Saf-4]
- Req-Saf-5: Control App must not send insulin bolus command when glucose is below 140 mg/dL or falling [UCA-Saf-5]
- Req-Saf-6: Control App must schedule insulin bolus to match expected post-meal glucose peak (within $\pm$30 minutes), based on glucose trend prediction algorithm [UCA-Saf-6]
- Req-Saf-7: Insulin Pump must deliver bolus insulin within 5 minutes of receiving command from Control App [UCA-Saf-8]
- Req-Saf-8: Insulin Pump must deliver only authorized insulin commands and reject invalid or manual override commands [UCA-Saf-9]
- Req-Saf-9: Insulin Pump basal infusion must not stop before completing programmed dose [UCA-Saf-11]
- Req-Saf-10: Insulin Pump basal infusion must not continue beyond programmed duration [UCA-Saf-10]
- Req-Saf-11: Control App must transmit patient data to Server during connectivity recovery, retrying every 5 minutes up to 3 attempts, to ensure remote monitoring [UCA-Saf-13]

- Req-Saf-12: Control App must validate data integrity before sending to Server, using checksum or hash verification [UCA-Saf-14]
- Req-Saf-13: Server must notify caregivers within 2 minutes about critical hypoglycemia (<70 mg/dL) or hyperglycemia (>300 mg/dL) events [UCA-Saf-15]
- Req-Saf-14: Server must avoid false alarms or misleading messages to caregivers, maintaining false-positive rate <5% [UCA-Saf-16]
- Req-Saf-15: Server notifications to caregivers must occur within a safe intervention time window of 2 minutes from event detection [UCA-Saf-17]

- Req-Sec-1: Insulin Pump must execute commands only from authenticated Control App using secure token-based authentication [UCA-Sec-1]
- Req-Sec-2: Control App must prevent unauthorized access or hijacking for insulin bolus commands, implementing multi-factor authentication [UCA-Sec-2]
- Req-Sec-3: CGM data transmission must be encrypted and protected against man-in-the-middle attacks using TLS 1.3 or equivalent [UCA-Sec-3]
- Req-Sec-4: Control App must prevent delayed or replayed bolus commands, rejecting any command older than 5 minutes [UCA-Sec-4]
- Req-Sec-5: Control App must prevent repeated bolus commands due to replay attacks using nonce or timestamp mechanisms [UCA-Sec-5]
- Req-Sec-6: Server must ensure integrity of alarms and notifications sent to caregivers using digital signatures or hash verification [UCA-Sec-6]
- Req-Sec-7: Control App must implement strong authentication for Pump access, including MFA and device binding [UCA-Sec-7]
- Req-Sec-8: Server must implement strict role-based access control for patient records [UCA-Sec-8]
- Req-Sec-9: Control App must encrypt all patient data transmitted to Server using AES-256 or equivalent [UCA-Sec-9]
- Req-Sec-10: Server must detect and prevent compromised infrastructure from sending delayed or manipulated feedback to caregivers, verifying message authenticity within 2 minutes [UCA-Sec-10]

Step 4: Identification of Loss Scenarios

Below are some examples of loss scenarios for the analyzed system.

Loss Scenarios: Safety

- LS-Saf-1: Due to a delayed or lost communication between the CGM and the Control App, the Control App does not issue the insulin bolus command when glucose levels rise above the threshold.

  - Related to: [UCA-Saf-1], [Req-Saf-1]
  - Leads to: Hyperglycemia [H-1] and potential harm to the patient [L-1].
- LS-Saf-2: The Control App issues an insulin bolus command during a period of normal or decreasing glucose levels, caused by an incorrect CGM reading.
  - Related to: [UCA-Saf-2], [Req-Saf-2]
  - Leads to: Hypoglycemia [H-1] and patient harm [L-1][L-2].

Loss Scenarios: Security

- LS-Sec-1: An attacker manipulates the CGM feedback to the Control App, sending falsified glucose values. The Control App then issues incorrect insulin commands to the Pump.
  - Related to: [UCA-Sec-1], [Req-Sec-1]
  - Leads to: Either lack of insulin delivery (hyperglycemia) or excessive insulin delivery (hypoglycemia) [T-1][T-2][L-1][L-2]
- LS-Sec-3: A denial-of-service (DoS) attack prevents the Control App from sending insulin delivery commands to the Pump in real time.
  - Related to: [UCA-Sec-4], [Req-Sec-4]
  - Leads to: Missed insulin delivery during hyperglycemia episodes [T-4] [L-1].

6.4 Chapter Summary

Chapter 6 presents the *STPA-SafeSecIoT* method, an extension of STPA designed to integrate the joint analysis and specification of safety and security requirements in critical IoT systems. The method is divided into four main steps: defining the purpose of the analysis, modeling the control structure, identifying unsafe control actions and specifying requirements, and describing loss scenarios.

We emphasize the importance of traceability between the analyzed elements and present a PoC applied to an automatic insulin delivery (AID) system, demonstrating the use of the proposed method in identifying critical safety and security requirements.

As contributions from *STPA-SafeSecIoT*, we highlight:

- A method for joint safety and security analysis, extending STPA, capable of mapping the emerging behaviors of complex systems resulting from the interaction between their subsystems/components;
- Support for the subsequent identification and classification of the relationships between safety and security requirements, as well as the resolution of possible conflicts between them;

- Support for the specification of safety and security requirements that reflects relevant information about these requirements and allows for their validation and use in support of the design, construction, and testing stages of critical IoT systems;
- Theoretical basis for the construction of a tool that supports the implementation of the stages of the proposed method in support of the analysis and specification of safety and security and, subsequently, the identification of relationships and conflicts between requirements.

CHAPTER 7

Evaluation of Critical IoT Systems Project Planning with Undergraduate Students

An essential step of the *Design Science Research* (DSR) methodology, adopted in this doctoral research, is the empirical evaluation of the developed artifacts. Considering the outcomes of this research, the *SafeSecIoT Canvas* is the initial element for carrying out the proposed RE process for safety and security of critical IoT systems, supporting project planning from its conception, as well as the definition of its scope, general system requirements, and the identification of IoT and security concerns. This chapter presents the preparation, execution, and analysis of the results of a controlled experiment and a survey conducted to evaluate the *SafeSecIoT Canvas* artifact.

7.1 Evaluation of the *SafeSecIoT Canvas* Artifact

STPA-based analysis is a process that requires not only specialized knowledge of this method for its application but also knowledge about the domain of the system to be analyzed, its objectives, and its scope. However, the STPA analysis [Leveson and Thomas, 2018], by default, begins with the definition of unacceptable losses for stakeholders, assuming that there is already prior planning of the system and that its scope has already been defined. In this regard, the *SafeSecIoT Canvas* artifact was proposed to fill the project planning gap and support the subsequent stages of the safety and security analysis and specification process.

The lack of project planning for a system, and of knowledge about its objectives, requirements, and other essential elements, can lead to misunderstandings in the comprehension of its scope and needs. Consequently, there is also a loss of purpose in the analysis and definition of safety and security requirements. In a critical IoT system, where these requirements must be ensured, such planning becomes essential to guarantee that the system is correctly defined and that safety and security requirements are properly understood and addressed from its very conception.

7.1.1 Experiment Design for Artifact Evaluation

To define the design of the experiment to be conducted, we considered the object to be evaluated (the developed artifact) and also the purpose of this object in relation to the research objective (why will it be used and what is the impact of its use?). In this case, the object to be evaluated in this experiment was the *SafeSecIoT Canvas* artifact, which was developed to support the safety and security RE process of critical IoT systems, with regard to the project planning.

To this end, we planned a controlled experiment with two groups from the same population, both performing the same task. One group used the artifact while the other did not, allowing us to analyze the impact of this treatment. In addition, we conducted a literature review to identify instruments suitable for evaluating different aspects of the task execution, with and without the artifact, as well as for analyzing both the task outputs and the artifact itself.

Based on these premises, Figure 7.1 presents the design that was defined for experimentation and evaluation of the *SafeSecIoT Canvas* artifact.

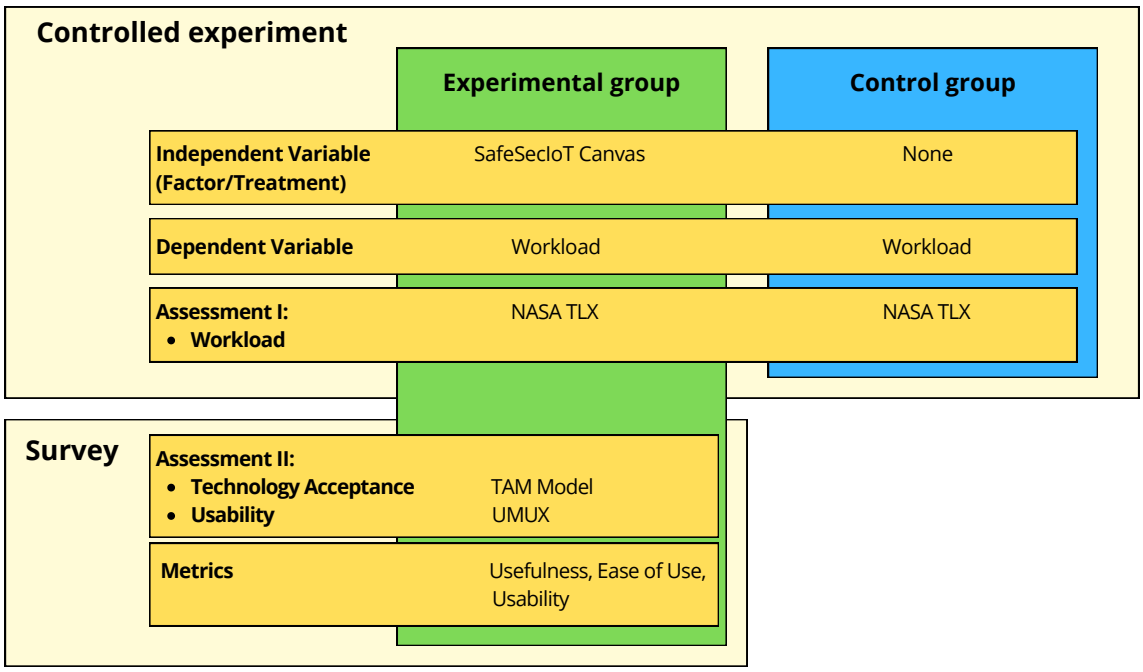


Figure 7.1: Planning the evaluation of the *SafeSecIoT Canvas* artifact.

This evaluation was divided into two parts. The first part, conducted through a controlled experiment with two groups (experimental and control), employed the *NASA Task Load Index* (NASA TLX) [Hart and Staveland, 1988] as the instrument to assess participants' workload during the task. The method considers participants' perceptions across six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. These aspects were evaluated after the completion of the task of

planning a critical IoT system project, in which the experimental group used the proposed artifact, while the control group did not and was free to define its own artifact. In addition to the quantitative evaluation, qualitative aspects were also examined through the analysis of the artifacts produced by the participants at the end of the experiment.

The second part of the evaluation was carried out exclusively with participants who used the *SafeSecIoT Canvas* artifact, with the objective of assessing its usefulness, ease of use, and overall usability. For this purpose, a survey was designed employing two well-established instruments: i) the *Technology Acceptance Model* (TAM) [Davis, 1989], and ii) the *Usability Metric for User Experience* (UMUX) [Finstad, 2010]. This approach made it possible to capture participants' perceptions regarding the artifact's quality and its perceived impact on task performance, as well as to analyze correlations between the quantitative results and the quality of the artifacts produced.

7.1.2 Using NASA TLX, TAM, and UMUX

As presented in the previous section, we combined the NASA-TLX, TAM, and UMUX methods to obtain a more comprehensive and complete evaluation of the *SafeSecIoT Canvas* artifact. This approach allowed us to evaluate not only the workload imposed on participants from different perspectives when using the developed artifact, but also its acceptance and perceived usability. Below, we detail how each of these instruments was used and how they complemented each other in this evaluation.

Each of these models (with their respective metrics) allows for the evaluation of different perspectives on user interaction with the system, which can help capture a more complete picture of the participants' experience:

- NASA-TLX: measures the cognitive workload perceived by a user when completing a task. The method considers six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. This allows us to assess the level of demand of the proposed artifact and whether it makes the task more difficult or easier than it would be without the use of this artifact, for example.
- TAM: evaluates the acceptance of technology based on perceived usefulness (PU) and perceived ease of use (PEOU), which directly influence the intention to use. This is essential to understand whether participants who used the *SafeSecIoT Canvas* artifact would consider adopting it in other projects for safety and security RE of IoT systems.
- UMUX: evaluates usability with a specific focus on simplicity and effectiveness of use, being a shorter and more focused metric than other usability models, such as SUS (*System Usability Scale*). UMUX is useful for measuring overall perception of usability quickly and efficiently.

7.2 Experiment Planning

7.2.1 Objective

The objective of the experiment was to evaluate whether the use of the *SafeSecIoT Canvas* artifact reduces the workload (NASA TLX) in the task of planning a critical IoT system project, compared to not using it, while delivering an artifact that maintains the quality of the project information. In addition, we evaluated the perception of participants who used the evaluated artifact regarding quality characteristics: usefulness and ease of use (TAM) and usability (UMUX). The evaluation was conducted with undergraduate students in the context of the Internet of Things course offered by the Bachelor's Degree in Artificial Intelligence (BIA) program at the Institute of Informatics (INF/UFG).

Below we present the detailed objective applying GQM:

- **Object of study:** to analyze the use of the *SafeSecIoT Canvas* artifact in the planning of a critical IoT system project;
- **Purpose:** to evaluate aspects of quality in use of the artifact, in comparison to its non-use, for the same project;
- **Quality focus:** with respect to workload [NASA,1986] and the quality of the artifact produced, and considering the usefulness and ease of use [Davis, 1989], in addition to the usability [Finstad, 2010] of the *SafeSecIoT Canvas* artifact, for participants who used it in the preparation of the project plan;
- **Point of view:** from the point of view of undergraduate students;
- **Environment:** in the context of a class on the IoT, offered by the Bachelor's Degree in Artificial Intelligence at INF/UFG.

7.2.2 Selection of Context and Participants

Population

The population defined for this experiment consisted of students enrolled in the IoT course at BIA/INF. This population was defined with the aim of establishing a prior level of knowledge among participants, taking advantage of the theoretical and practical training provided during the course and the specific knowledge necessary for the development of an IoT system, which students would have at the end of the course. Prior knowledge about this type of system was an important requirement for defining the population, since it is essential for carrying out the experimental activity to be performed and could negatively impact the results of the study if not met.

Sample and selection of participants

The sample considered for the experiment consisted of students regularly enrolled in the IoT course in the 2024/2. Participation was voluntary, upon invitation and signing of an informed consent form, in accordance with the research project[1] approved by the Ethics and Research Committee (CEP/UFG). Under these terms, we had a total of 23 students participating in the experiment. The evaluation was carried out with a single class, since the IoT course is offered only once a year, in the second semester.

Definition of groups

To conduct the experiment, the 23 participants were randomly assigned to two groups:

- Experimental group: which received the treatment in the experiment, in this case the use of the *SafeSecIoT Canvas* artifact to perform the task of planning the design of a critical IoT system;
- Control group: which performed the same task but did not receive any type of treatment.

It was defined in the experiment design that the groups (experimental and control) should have the same number of participants. However, due to the odd number of total participants, 12 students were allocated to the experimental group (divided into 4 teams of 3 participants) and 11 students to the control group (divided into 3 teams of 3 participants and one team of 2 participants).

7.2.3 Formulation of Hypotheses

Based on the objective proposed for the evaluation, presented in subsection 7.2.1, the following hypotheses were defined for the experiment:

- H01 – There will be no significant difference in the perceived workload between the experimental group (which performed the task using the *SafeSecIoT Canvas*) and the control group (which performed the same task without using the artifact).
  - HA01_A – The use of the *SafeSecIoT Canvas* artifact contributes to reducing the workload in the task of developing a project plan for a critical IoT system compared to not using it.

[1] Registered under the Certificate of Ethical Review: 79204024.8.0000.5083

  - HA01_B – The use of the *SafeSecIoT Canvas* artifact contributes to an increase in the workload in the task of developing a project plan for a critical IoT system compared to not using it.

- H02 – The *SafeSecIoT Canvas* artifact is not well accepted in use in supporting the task of developing a project plan for a critical IoT system.
  - HA02 or ¬H02 (negation of hypothesis H02): the *SafeSecIoT Canvas* artifact is well accepted in use in supporting the task of developing a project plan for a critical IoT system.

- H03 - The *SafeSecIoT Canvas* artifact does not have good usability in supporting the task of developing a project plan for a critical IoT system.
  - HA03 = ¬H03 (negation of hypothesis H03): the *SafeSecIoT Canvas* artifact has good usability in supporting the task of developing a project plan for a critical IoT system.

7.2.4 Instrumentation

As presented in subsections 7.1.1 and 7.1.2, the instruments defined for conducting the controlled experiment and the survey were the NASA TLX, the TAM, and the UMUX. Below we briefly present how these methods and their metrics were used in the context of this evaluation.

Variable selection

For the controlled experiment, the independent variable (also called the main factor of the experiment) is the "type of artifact" to be used for developing the critical IoT system design plan. This variable can take two distinct values: i) *SafeSecIoT Canvas* artifact (for the experimental group) and ii) none (for the control group). Meanwhile, the dependent variable will be the "workload." The objective is to measure how the independent variable (type of artifact used) influences the dependent variable (workload).

Thus, the control group will provide a standard, or reference, that will allow us to assess whether the treatment performed in the experimental group (use of the *SafeSecIoT Canvas* artifact) has an effect (positive or negative) on the required workload and on the final product of the experiment. This comparison between the results of the groups will be used to confirm or refute the hypotheses presented in subsection 7.2.3.

Usefulness, ease of use, and usability can also be considered dependent variables. However, these metrics will not be compared between groups. Since no treatment was applied to the control group, these variables will only be used to evaluate the *SafeSecIoT*

Canvas artifact, and the data will be collected through a questionnaire applied to the experimental group.

Questionnaires administered

Three types of questionnaires were administered after the experimental activity: i) a questionnaire based on NASA TLX to assess perceived workload (administered to both groups); ii) a questionnaire based on TAM and UMUX, to assess the *SafeSecIoT Canvas* artifact (only for the experimental group); and iii) a qualitative questionnaire with open-ended questions (specific to each group).

To assess task workload, the NASA TLX questionnaire was applied. This instrument uses a scale to evaluate each aspect individually, followed by pairwise comparisons to determine the relative relevance of the dimensions. The workload questionnaire was made available online to participants in both the experimental and control groups.

To evaluate the PU, PEOU, and usability of the artifact, two additional instruments were employed. The *Technology Acceptance Model* (TAM) [Davis, 1989] was used, consisting of six questions addressing usefulness and six addressing ease of use. In addition, the UMUX questionnaire [Finstad, 2010], composed of four questions, was applied to assess usability. Both instruments used a 7-point Likert scale. This evaluation questionnaire was made available online exclusively to participants in the experimental group, who interacted with the *SafeSecIoT Canvas* artifact. Further details on the instruments and their application can be found in the Appendix A.

7.3 Data Analysis and Interpretation of Results

7.3.1 Descriptive Statistics

After conducting the experiment and collecting data, as presented in Appendix A (Section A.1), the information was organized, allowing the presentation of essential information in the form of tables and graphs. In addition, calculations of the main representative parameters were performed using descriptive statistics.

Based on this understanding of each participant's assessment of the workload, we can analyze the statistical data presented in Table 7.1. These descriptive data were obtained from the processing of the database collected (from participants) during the experiment, which was prepared and subsequently processed using the Jamovi tool. Initially, we present a summary of the descriptive data regarding the participants' responses to the NASA TLX questionnaire. Next, we discuss the main findings in relation to the parameters analyzed.

Table 7.1: Descriptive statistics: ratings obtained from the NASA TLX questionnaire.

	Group	**Work-load**	**Mental**	**Physical**	**Tempo-ral**	**Perfor-mance**	**Effort**	**Frustra-tion**
N	Control	11	11	11	11	11	11	11
	Experimental	12	12	12	12	12	12	12
Missing	Control	0	0	0	0	0	0	0
	Experimental	0	0	0	0	0	0	0
Mean	Control	61.4	61.4	8.18	70.9	57.3	50.9	47.3
	Experimental	48.9	62.1	14.6	49.2	36.3	58.3	25.8
95% CI	Control	49.8	46.4	2.12	58.4	44.9	35.1	26.1
(lower)	Experimental	41.4	52.2	9.43	38.6	18.6	48.2	16.3
95% CI	Control	73.0	76.3	14.2	83.4	69.7	66.7	68.5
(upper)	Experimental	56.3	72.0	19.7	59.7	53.9	68.5	35.4
Median	Control	56.0	70	5	75	55	55	50
	Experimental	45.2	62.5	15.0	52.5	25.0	55.0	25.0
SD	Control	17.3	22.3	9.02	18.5	18.5	23.5	31.6
	Experimental	11.7	15.6	8.11	16.6	27.7	16.0	15.1
Variance	Control	299	495	81.4	344	342	554	997
	Experimental	136	243	65.7	277	769	256	227
Minimum	Control	35.0	30	5	45	30	5	5
	Experimental	37.0	35	5	20	5	35	5
Maximum	Control	86.3	90	35	100	90	80	100
	Experimental	71.3	90	30	85	85	85	50
Shapiro-Wilk W	Control	0.924	0.906	0.419	0.933	0.977	0.942	0.952
	Experimental	0.874	0.976	0.826	0.932	0.876	0.953	0.936
Shapiro-Wilk p	Control	0.356	0.216	$<.001$	0.440	0.949	0.544	0.674
	Experimental	0.073	0.961	0.019	0.404	0.077	0.682	0.443

Note. SD = Standard Deviation. The 95% confidence interval for the mean assumes the sampling distribution follows a t distribution with $N-1$ degrees of freedom.

The experiment involved 23 participants, who were divided into two groups: experimental (12 participants) and control (11 participants). All participants completed the questionnaire in its entirety, with no omissions in any of the mandatory questions.

Measures of Central Tendency

We can observe that the average workload perceived by the control group was higher than that of the experimental group (61.4 and 48.9, respectively). This number is an initial indication that the treatment applied to the experimental group (the use of the *SafeSecIoT Canvas* artifact) may have contributed to reducing the workload for performing the task. This indication needs to be confirmed (or refuted) through statistical tests (presented later in this chapter) so that we can confirm or refute hypothesis H01.

Still on the averages, analyzing the parameters evaluated, we can observe that the "Mental Demand" was very similar between the groups (a difference of only 0.7 percentage points more for the experimental group). On the other hand, the evaluations of "Temporal Demand" and "Frustration" were significantly lower (i.e., better) for the experimental group (differences of 21.7 and 21.5 percentage points, respectively, for the control group). Along these lines, "Performance" was significantly better for the experimental group (a difference of 21 percentage points compared to the control group). Finally, the averages for "Physical Demand" and "Effort" were slightly higher for the experimental group (differences of 6.42 and 7.4 percentage points, respectively). This assessment is quite understandable, since the artifact used by the experimental group (treatment) was a printed canvas, which requires greater direct physical interaction, both from the participants with the artifact (to fill it out) and among themselves.

Measures of Dispersion

Measures of dispersion are important for a complete analysis of the data, allowing the evaluation of how much the data varies around the measures of central tendency. Thus, it is the measures of dispersion that show whether the measured measures of central tendency are representative or not for the analyzed data set.

A first measure that can be analyzed in this case is amplitude. We can observe that, for the experimental group, the amplitude for the workload was 34.3 (with a maximum value of 71.3 and a minimum of 37), while for the control group it was 51.3 (maximum of 86.3 and minimum of 35). These numbers show that the data set (workload values) obtained from the participants in the experimental group was more homogeneous than the data from the control group, as they were distributed within a significantly smaller range. In other words, the participants in the experimental group evaluated the workload in a more similar or closer manner to each other.

Considering the scores assigned to the evaluation parameters by the participants, the only parameter that had a higher maximum value for a participant in the experimental group was "Effort," with the maximum values for the other parameters found in the control group's evaluations. Similarly, the only parameter that had a higher standard deviation in

the control group was “Performance,” with lower values in all comparisons for the other parameters.

Finally, we calculated the coefficient of variation of the data, that is, the percentage of variability of the collected data. This percentage is calculated based on the standard deviation (SD) and the mean (M), where: (SD*100)/M. Thus, we have that:

- For the experimental group, the coefficient of variation of the data in relation to the mean was: (11.7*100)/48.9 = 23.92%.
- For the control group, the coefficient of variation of the data was: (17.3*100)/61.4 = 28.17%.

The values obtained show good consistency between the data and, as will be shown in later graphs, the absence of outliers.

Graphical Analysis

Figure 7.2 shows the boxplot graph of workload assessments for the control and experimental groups.

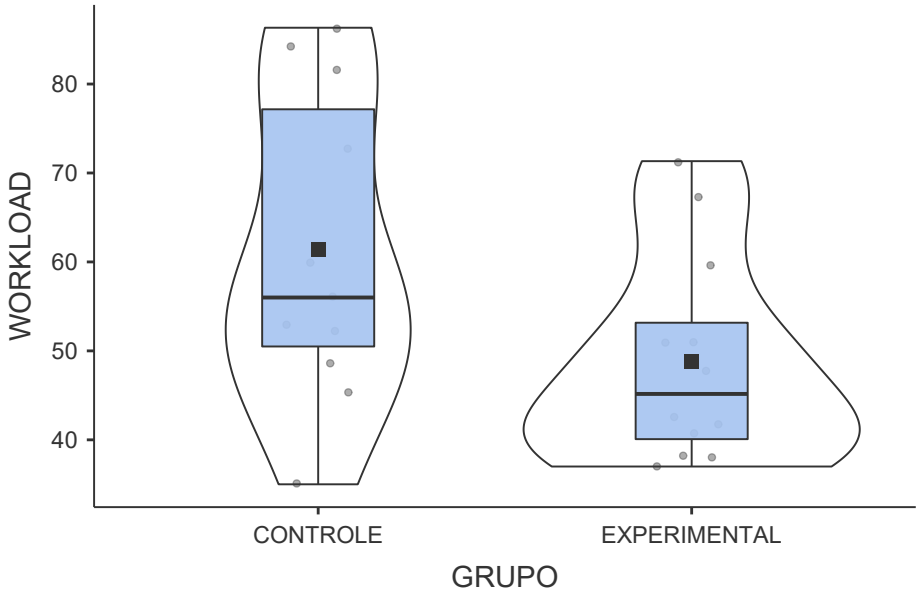


Figure 7.2: Box plot graph for workload: experimental and control.

This graph shows the data (separated by groups, where N=12 for the experimental group and N=11 for the control group), divided into three quartiles:

- 1st quartile (line at the bottom): indicates where the lowest 25% of the collected workload values are located (in this case, 3 values);

- 2nd quartile (closed box between the lines): indicates where the 50% central workload values collected are located (in this case, 6 values for the experimental group and 5 for the control group); and
- 3rd quartile (line at the top): indicates where the 25% highest workload values collected for the group are located (in this case, 3 values).

Regarding the distribution of workload values obtained by participants in each group (represented by the points in the graph), we can state that there were no outliers (values outside the normal range). All values fell within the three quartiles. One point to note, as previously mentioned, is that the range for workload values is greater in the control group than in the experimental group, which indicates greater homogeneity among the values in the latter. Considering the second and third quartiles, this range is even smaller in the experimental group, as evidenced by the violin added to the graph, showing that the lowest workload value is much closer to the median than the highest value in this group, which is also reflected in the group's average workload.

Correlation Analysis

Correlation analysis is a statistical technique used to measure the strength and direction of the relationship between two quantitative variables. It is widely used to understand how changes in one variable are associated with changes in another. We analyzed the data from the controlled experiment to understand the correlation between the NASA TLX metrics, as assessed by the participants, and the calculated workload. To do this, we analyzed the data from the experimental and control groups separately.

Figure 7.3 shows the correlation matrix for the experimental group.

CORRELATION MATRIX		WORKLOAD	MENTAL DEMAND	PHYSICAL DEMAND	TEMPORAL DEMAND	PERFORMANCE	EFFORT	FRUSTRATION
MENTAL DEMAND	R de Pearson	0.439	—					
	p-value	0.153	—					
PHYSICAL DEMAND	R de Pearson	-0.010	-0.370	—				
	p-value	0.975	0.236	—				
TEMPORAL DEMAND	R de Pearson	0.504	0.376	-0.037	—			
	p-value	0.095	0.229	0.910	—			
PERFORMANCE	R de Pearson	0.681	-0.227	0.417	0.002	—		
	p-value	0.015	0.477	0.177	0.994	—		
EFFORT	R de Pearson	0.726	0.307	-0.181	0.396	0.359	—	
	p-value	0.008	0.332	0.573	0.203	0.252	—	
FRUSTRATION	R de Pearson	0.120	-0.396	0.115	0.148	0.253	0.214	—
	p-value	0.711	0.203	0.722	0.645	0.427	0.504	—

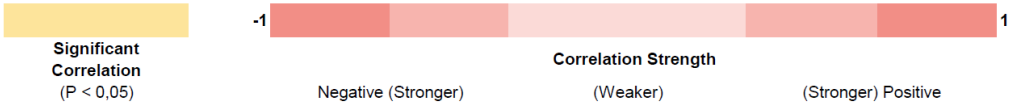


Figure 7.3: Correlation matrix between workload and NASA TLX metrics for the experimental group.

It was observed that only the “Performance” and “Effort” metrics showed a strong correlation with the overall workload. In contrast, “Mental Demand” and “Temporal Demand” presented moderate correlations, while “Physical Demand” and “Frustration” showed virtually no correlation with the main measure.

Analyzing the participants’ evaluation data, we observed that this behavior occurs because the evaluations for the “Performance” and “Effort” metrics are the ones that most closely approximated the workload values for the participants in the experimental group. Meanwhile, the other metrics had greater variation and therefore do not correlate significantly with the workload.

Still considering the experimental group, the six metrics that make up the NASA TLX did not have significant correlations with each other. In other words, it is not possible to say that the variation (increase or decrease) in the evaluation in relation to one metric significantly influences another.

Figure 7.4 shows the correlation matrix for the control group. We can observe that, for this group, there are a greater number of significant correlations between workload and the metrics analyzed, and furthermore, they are stronger. In this case, the metrics “Mental Demand,” “Temporal Demand,” “Performance,” “Effort,” and “Frustration” had a significant correlation with the calculated workload.

CORRELATION MATRIX		WORKLOAD	MENTAL DEMAND	PHYSICAL DEMAND	TEMPORAL DEMAND	PERFORMANCE	EFFORT	FRUSTRATION
MENTAL DEMAND	R de Pearson	0.847	—					
	p-value	< .001	—					
PHYSICAL DEMAND	R de Pearson	-0.236	-0.323	—				
	p-value	0.486	0.333	—				
TEMPORAL DEMAND	R de Pearson	0.677	0.251	0.056	—			
	p-value	0.022	0.457	0.871	—			
PERFORMANCE	R de Pearson	0.813	0.569	-0.467	0.693	—		
	p-value	0.002	0.068	0.147	0.018	—		
EFFORT	R de Pearson	0.676	0.713	-0.298	0.135	0.385	—	
	p-value	0.022	0.014	0.374	0.692	0.242	—	
FRUSTRATION	R de Pearson	0.928	0.778	-0.397	0.636	0.885	0.481	—
	p-value	< .001	0.005	0.227	0.035	< .001	0.134	—

Significant Correlation (P < 0,05)

-1 Correlation Strength 1

Negative (Stronger) (Weaker) (Stronger) Positive

Figure 7.4: Correlation matrix between workload and NASA TLX metrics for the control group.

Analyzing the participants’ evaluation data, we can affirm that this correlation exists in the control group because there was consistency in higher evaluations for all metrics, with the exception of “Physical Demand.” This directly impacted a higher workload for the Control group, which correlates with the metrics analyzed.

In addition, in the control group, some metrics had significant correlations with each other. For example:

- "Temporal Demand" and "Performance": an increase in temporal demand impacts performance negatively (and vice versa, poor performance is linked to temporal demand for task completion);
- "Time Demand" and "Frustration": an increase in time demand impacts an increase in frustration with the task (and vice versa, frustration is linked to time demand for completing the task);
- "Mental Demand" and "Effort": mental demand impacts the increase in perceived effort to perform the task (and vice versa, increased effort is linked to high mental demand);
- "Mental Demand" and "Frustration": high mental demand is linked to increased frustration in performing the task (and vice versa, high frustration is linked to high mental demand);
- "Performance'' and "Frustration": low performance in completing the task impacts an increase in frustration (and vice versa, a high level of frustration is linked to low performance).

7.3.2 Inferential Statistics

In this subsection, we apply inferential statistics to analyze data from the collected sample to verify whether it is possible to make generalizations or draw conclusions about the population. Thus, we statistically analyze whether the sample we collected is representative and whether it is possible to infer information about the rest of the population based on this sample.

Before applying a statistical test to the data, it is necessary to understand what type of test applies to the collected data and the type of study conducted. In this case, the overall objective of the study is to compare two groups (experimental and control) with different types of treatment (use of the *SafeSecIoT Canvas* artifact and non-use).

Analyzing the groups involved, we can say that they are independent (or unpaired) groups. In other words, they are groups formed by different participants, but from the same population (students of the IoT course). In addition, another parameter that needs to be defined is the type of variable analyzed. In this case, we are evaluating the workload, which is a quantitative variable.

Finally, it is essential to understand the type of distribution of the sample data (whether it is normal or not) so that the correct test for the type of distribution can be applied. For this, the normality test is applied. The Jamovi tool, which we use for data analysis, allows the application of the most common normality test, which is the Shapiro-

Wilk test. This test results in a p-value, which indicates whether the distribution is normal (p>0.05) and in this case (based on the experiment data) the Student's t-test could be applied, or whether the distribution is considered different from normal ($p < 0.05$), in which case the Mann-Whitney test could be applied.

The result of the normality test for workload was presented alongside the descriptive statistical data in Table 7.1. The calculated Shapiro-Wilk p-value was:

- Control group: 0.356;
- Experimental group: 0.073.

That is, for both groups, the data distribution was considered normal. In this case, we applied Student's t-test to verify whether it is possible to generalize the sample data to the population. The test was applied using the Jamovi tool, and the result is presented in Table 7.2:

Table 7.2: Results of the Student's t-test.

Variable	**Test**	**Statistic**	**df**	*p*
Workload	Student's *t*	2.0543	21.0	0.053

Based on the t-test, two hypotheses can be evaluated:

- Null hypothesis (H0): There is no significant difference between the groups; and
- Alternative hypothesis (HA): There is a significant difference between the groups.

The general rule for evaluating test results is as follows:

- If $p \leq \alpha$ (usually α=0,05)): we reject H0, i.e., the groups are statistically different;
- If $p > \alpha$: we do not reject H0, i.e., there is insufficient evidence to claim that the means are different.

In the test performed, the value p = 0.053. This value is at the limit of α for us to reject the hypothesis H0 and consider the groups statistically different. This value means that the error rate in the classification of other samples from the population, based on the data from this experiment, would be 5.3%.

To complement this test, we can analyze the descriptive graph that shows the confidence intervals (CI) for the experimental and control groups. CIs are used to indicate the reliability of an estimate. The graph in question is shown in Figure 7.5:

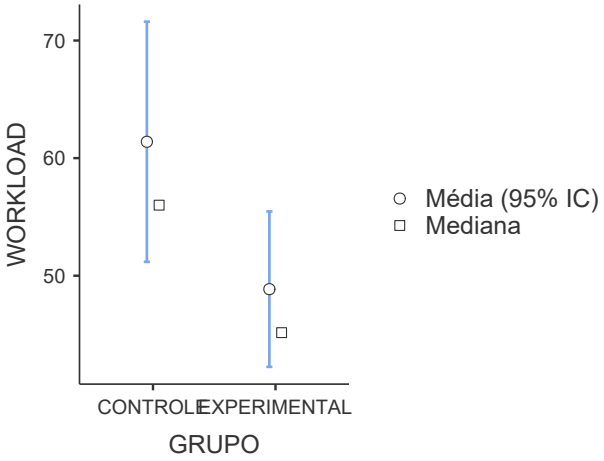


Figure 7.5: Descriptive graph.

In this case, we see that the confidence interval of the control group does not overlap with the mean of the experimental group. In other words, we can say that there is a statistical difference between the groups, rejecting the null hypothesis (H0).

Discussion of results for workload

The average workload for the experimental group was 24.74% lower than that of the control group, indicating that the treatment applied (the use of the *SafeSecIoT Canvas* artifact) contributed to reducing the workload for planning the design of a critical IoT system. In addition, it was observed that the use of the artifact had a positive impact on the perception of lower time demand and frustration, as well as a better assessment of task performance. The mental demand observed was very similar between the groups, and although the physical and effort demand indices showed higher average values for the experimental group, they were offset by the significant difference in the other items and, consequently, in the final average workload.

As demonstrated, based on the analysis of the results obtained by applying descriptive and inferential statistics, we can affirm that "the use of the *SafeSecIoT Canvas* device contributes to reducing the workload in the task of developing a project plan for a critical IoT system compared to not using it", thus **refuting hypothesis H01 presented at the beginning of this chapter and validating hypothesis HA01_A**.

7.3.3 Quantitative and qualitative analysis of artifacts

In addition to understanding whether the use of the *SafeSecIoT Canvas* artifact reduces the perceived workload in the planning activity for critical IoT system projects, an analysis was also conducted to understand whether the quality of the artifacts produced is compatible with what is expected for the task performed. Understanding the quality

of the results produced by the participants can provide valuable information about the effectiveness of using *SafeSecIoT Canvas*, including the accuracy and completeness of the results obtained.

To analyze the results produced by the participants in the experimental and control groups in a way that minimizes the introduction of bias, we used the Iramuteq tool[2] to analyze the documents produced in the experiment. Iramuteq (or Interface de R *pour les Analyses Multidimensionnelles de Textes et de Questionnaires*) is a statistical tool that allows data analysis in qualitative research with the integration of quantitative levels, bringing greater objectivity and advances to the interpretation of text data.

Preparation of Data for Analysis

In order to perform the quantitative-qualitative analysis using the Iramuteq tool, the project planning information generated by the groups must be transcribed into a corpus. For this reason, two corpora were generated from the experiment (which we call the experimental corpus and the control corpus), one for each group of participants. In addition, a third corpus was generated, called the reference corpus, with reference data for the project planning of the Automatic Insulin Delivery system, produced by the researchers. The reference corpus will be considered, in the context of this analysis, as a "template" or reference for the other corpora.

Figure 7.6 shows an excerpt from the corpus prepared for data analysis:

```
**** *corpus_01 *grupo_00
As Justificativas para o Projeto são: dificuldades no autocuidado de pessoas com diabetes tipo 1; aferição manual dos níveis de
glicose; complexidade de decisão e controle manual sobre aplicação de insulina; desgaste físico e mental de pessoas com diabetes
tipo 1; altas variações da taxa de glicose prejudicando a qualidade de vida do portador de diabetes tipo 1. O Objetivo é: desenvolver
um sistema IoT (AID) para monitoramento de glicemia em tempo real e entrega automática de insulina no organismo. Os Benefícios
são: aferição automática dos níveis de glicose; controle automatizado do nível de glicose e de entrega e de insulina; menor variação
nos níveis de glicose no sangue (AID simulando a função de um pâncreas saudável); ...
**** *corpus_02 *grupo_01
**** *corpus_02 *grupo_02
**** *corpus_02 *grupo_03
**** *corpus_02 *grupo_04
**** *corpus_03 *grupo_05
**** *corpus_03 *grupo_06 ...
**** *corpus_03 *grupo_07 ...
**** *corpus_03 *grupo_08 ...
```

Figure 7.6: Excerpt from the corpus (prepared for use).

The transcription for each group is described as a distinct text (group) within the corpus. These groups were separated into three distinct corpora:

- corpus_01 (group_00): this is the reference corpus (or template), whose data was produced by the researchers before the experiment was conducted;

[2] http://www.iramuteq.org/

- corpus_02 (group_01, group_02, group_03, and group_04): this is the data produced by the teams in the experimental group;
- corpus_03 (group_05, group_06, group_07, and group_08): these are the data produced by the control group teams.

From the prepared corpus, the data produced in the context of the experiment can be analyzed and compared with the reference data for the project planning document that was expected as a result. Below, we present the main analyses obtained from processing this corpus in Iramuteq.

Correspondence Factor Analysis (CFA)

Correspondence Factor Analysis (CFA) is a statistical technique that explores the association between categories of variables in a contingency table. In Iramuteq, it is used to relate words/text segments to defined categories (e.g., groups of respondents, variables, response classes, etc.). The goal is to analyze and graphically represent the relationships found between the analyzed corpora.

Figure 7.7 shows an CFA graph generated by the Iramuteq tool. This graph allows for the interpretation of the distribution and general proximity of the groups, as well as the relationships between the reference, experimental, and control corpora.

The reference corpus (X.group_00) is located in the lower right, relatively close to some groups in the Experimental corpus (X.group_01, X.group_02, and X.group_03). This result indicates that these groups in the experimental corpus are more similar to the reference corpus. The last group of the experimental corpus (X.group_04) begins to distance itself from the Reference corpus and the other groups of the Experimental corpus, indicating that there is some heterogeneity within the Experimental group itself.

The groups in the control corpus (X.group_05, X.group_06, X.group_07, and X.group_08) are more dispersed and positioned in opposite regions of the graph in relation to the reference corpus. This result indicates less semantic similarity between the groups in the control corpus and the reference corpus. In addition, the Control groups are more dispersed, and one of the groups (X.group_06) is quite distant from both the groups in the same corpus and the others, suggesting that it has significantly different characteristics.

Interpreting the dimensions of the graph, we can analyze that:

- Dimension 1 (27.6%): Explains most of the variability in the graph. The group X.group_06 from the control corpus is well removed from the reference corpus and also from the other groups in the same corpus, suggesting very distinct characteristics. X.group_02 (experimental), on the other hand, is the most aligned with the Reference corpus on this axis.

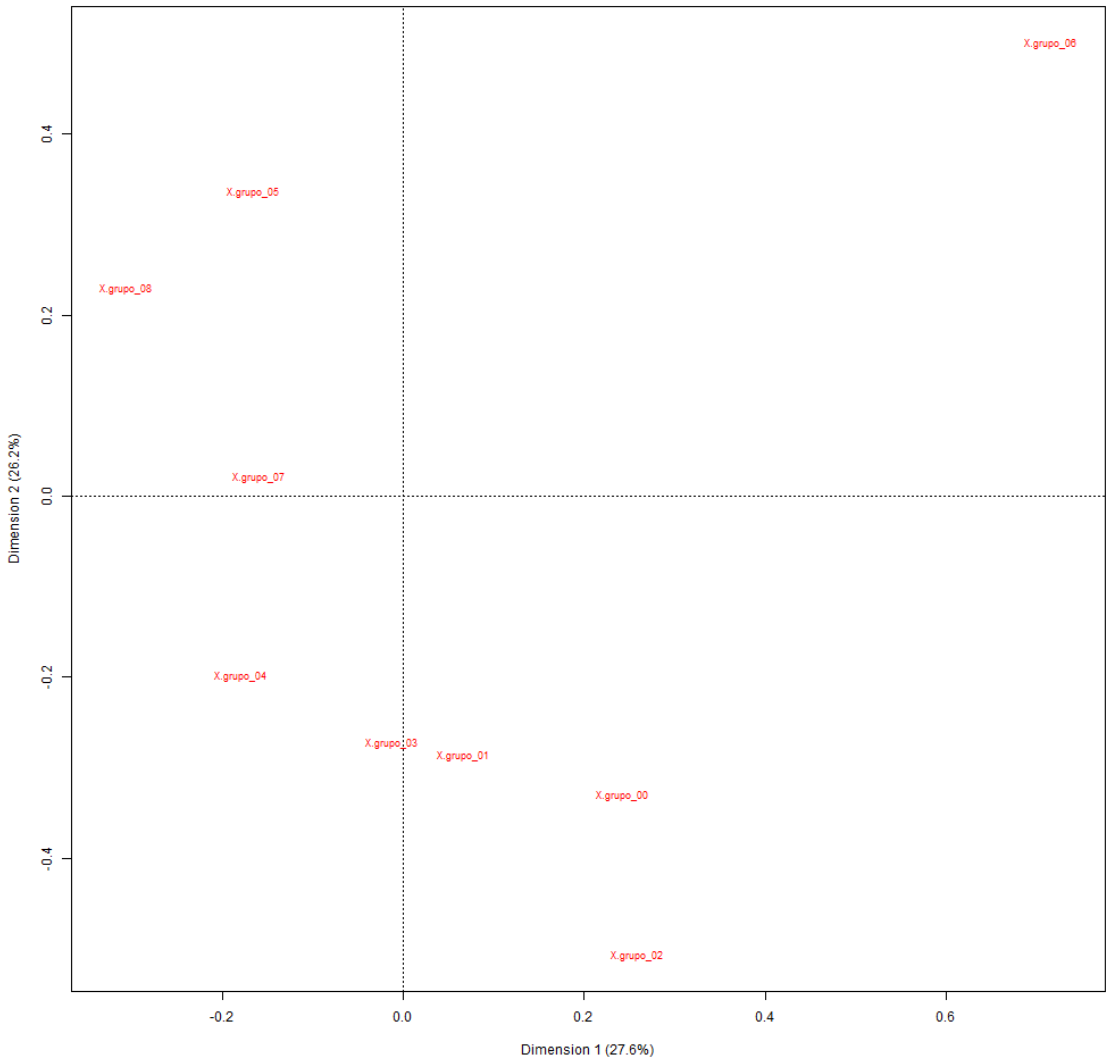


Figure 7.7: Correspondence Factor Analysis chart for the corpora.

- Dimension 2 (26.2%): Shows secondary differences between the corpora. The groups in the control corpus are distributed in a higher position, while the groups in the experimental corpus and the reference corpus occupy positions at the bottom of this axis. This reinforces that the artifacts produced by the Control groups have characteristics that diverge from the other artifacts (from the Experimental and Reference groups) from a semantic point of view.

In the CFA graph generated by Iramuteq, each dimension (or axis) is accompanied by a percentage, cited in the previous analysis. This value represents what the percentage signifies, that is, the proportion of variance (or inertia) in the data explained by that axis. In other words, the percentage indicates how much of the information about the relationships between categories and words is represented in that dimension.

Conclusions about the analyzed corpora:

- Experimental:
  - Some groups (X.group_01, X.group_02, and X.group_03) are closer to the Reference, indicating greater adherence to the expected semantics.
  - One group (X.group_04) is further away, suggesting less proximity, but is still closer than the groups in the control corpus.
- Control:
  - It is considerably distant from the Reference corpus, with dispersion among the group's own teams (X.group_05, X.group_06, X.group_07, X.group_08), indicating different data among the results produced by the control group. This indicates low similarity with the template and possible misalignment with the central theme.

Labbé Distance Matrix

The Labbé distance matrix represents the similarity or proximity between groups based on CFA. Each value in the matrix reflects the distance between two groups, where lower values indicate greater similarity between the analyzed contents (considering the corpus of each group). Figure 7.8 shows the distance matrix considering the groups participating in the controlled experiment and the reference corpus/group.

Regarding the interpretation of the matrix, lighter colors (close to light orange) indicate greater distance between groups, i.e., less similarity. The darker colors (close to bright orange or red) indicate a smaller distance between the groups, i.e., greater similarity. The diagonal will always have a value of zero, as it represents the distance of a group from itself, and will be white.

Regarding the values, which represent distance, intersections with values close to 0 suggest that the groups share much of the vocabulary or have semantically similar texts. Higher values suggest that the groups address different topics or use different vocabularies.

The results reinforce that the experimental groups were closer to each other (and to the reference group), while the control groups showed a greater distance overall. This finding reinforces that the use of the *SafeSecIoT Canvas* artifact provided greater consistency and effectiveness in terms of the results obtained, when compared to work carried out without a model or other type of artifact used for project planning.

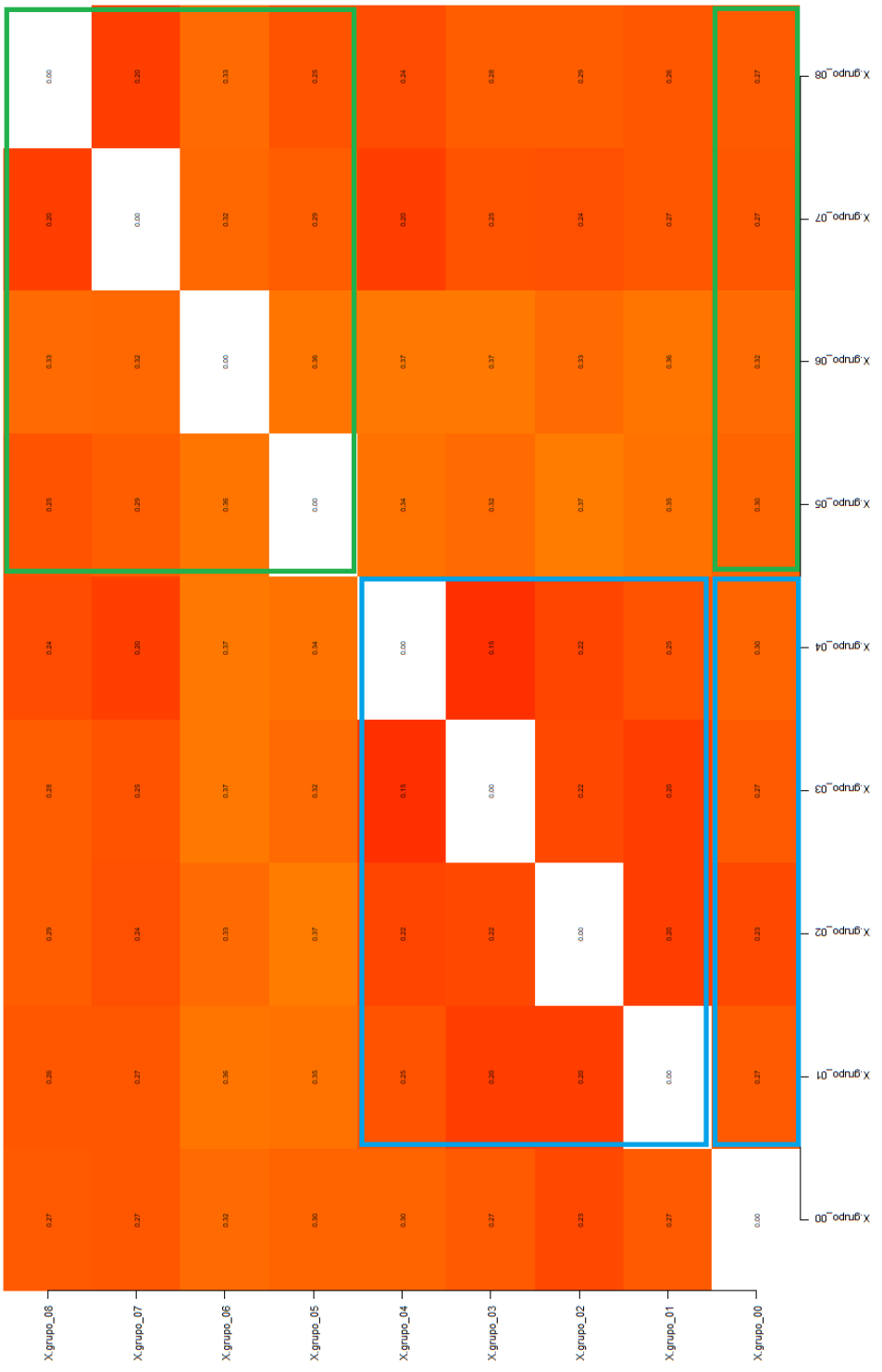


Figure 7.8: Labbé Distance Matrix for the analyzed corpora.

7.4 Survey with Participants of the Experimental Group

In addition to applying the NASA TLX to the control and experimental groups, a survey based on TAM and UMUX was also applied to the experimental group in order to evaluate the quality in use of the *SafeSecIoT Canvas* artifact.

7.4.1 Reliability of Questionnaires

To ensure the reliability of the questionnaires used in the evaluation of the *SafeSecIoT Canvas* artifact, Cronbach's alpha coefficient was used. This procedure verifies the internal consistency of items organized into theoretical constructs, such as PU, PEOU, and UMUX, by analyzing the homogeneity of responses. Thus, the use of Cronbach's alpha reinforces the statistical validity of the research and ensures that the instruments used are adequate to capture the experts' perceptions of the acceptance and adequacy of the proposed solutions. Table 7.3 shows Cronbach's alpha, mean, and standard deviation for each of the constructs used in the survey for the experimental group.

Table 7.3: Summary metrics by construct (mean, SD, and Cronbach's α)

Construct	**Mean**	**Standard deviation**	**Cronbach's α**
Perceived Usefulness (PU)	6.03	0.873	0.895
Perceived Ease of Use (PEOU)	5.82	0.978	0.941
Usability (UMUX)	5.67	0.979	0.779

The results indicate that the instruments used showed high statistical reliability, lending legitimacy to the analyses based on the perceptions of the participants in the experimental group. The high averages (all above 5.67 on a scale of 1 to 7) reinforce the positive evaluation of the artifact. Further details will be explored in the next section with the support of other descriptive statistics.

7.4.2 Analysis and Interpretation of Results

Perceived Usefulness (PU-TAM) of the *SafeSecIoT Canvas* artifact

Table 7.4 presents the summary of descriptive statistics for PU of *SafeSecIoT Canvas*. It can be observed that the high mean PU (6.03) and relatively low standard deviation (0.873) are the result of a positive and consistent evaluation of all items in the construct by the participants, with the means of all items ranging from 5.83 (the lowest) to 6.25 (the highest).

Table 7.4: Descriptive statistics for PU of the *SafeSecIoT Canvas*.

	PU-TAM-1	**PU-TAM-2**	**PU-TAM-3**	**PU-TAM-4**	**PU-TAM-5**	**PU-TAM-6**
N	12	12	12	12	12	12
Missing	0	0	0	0	0	0
Mean	6.00	5.83	6.08	6.08	5.92	6.25
Median	6.00	6.00	6.50	6.00	6.00	7.00
Standard deviation	0.953	0.937	1.08	1.08	1.16	1.22
Minimum	4	4	4	4	3	3
Maximum	7	7	7	7	7	7

Perceived Ease of Use (PEOU-TAM) of the *SafeSecIoT Canvas* artifact

Regarding PEOU, the overall mean of the construct (5.82) and its items was slightly lower than that of PU, but still considered high (on a scale of 1 to 7). The highest average among the items was 5.92 and the lowest was 5.67, showing high consistency among the evaluations and high reliability for the questionnaire (alpha = 844). The other descriptive statistics are presented in Table 7.5.

Table 7.5: Descriptive statistics for PEOU of the *SafeSecIoT Canvas*.

	PEOU-TAM-1	**PEOU-TAM-2**	**PEOU-TAM-3**	**PEOU-TAM-4**	**PEOU-TAM-5**	**PEOU-TAM-6**
N	12	12	12	12	12	12
Missing	0	0	0	0	0	0
Mean	5.67	5.92	5.92	5.67	5.92	5.83
Median	5.50	6.00	6.00	5.50	6.50	6.00
Standard deviation	1.15	1.08	0.900	1.15	1.38	0.937
Minimum	4	4	4	4	3	4
Maximum	7	7	7	7	7	7

Usability (UMUX) of the *SafeSecIoT Canvas* artifact

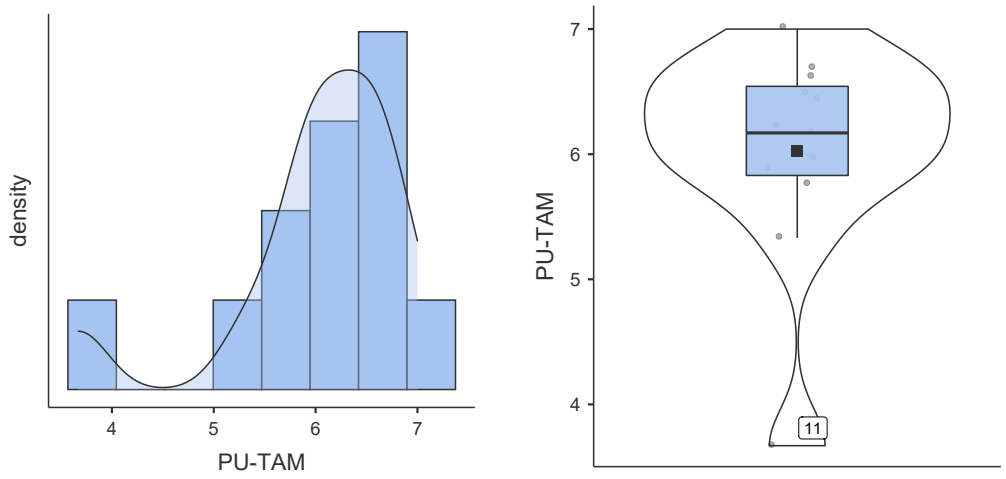
In addition to technology acceptance, based on TAM constructs, we also used UMUX to evaluate the usability of *SafeSecIoT Canvas*. The overall average for the construct (5.67) was slightly lower than that of the other constructs analyzed, but it shows that the responses were predominantly positive. The other descriptive statistics are presented in Table 7.6.

Table 7.6: Descriptive statistics for usability of *SafeSecIoT Canvas*.

	UMUX-1	**UMUX-2**	**UMUX-3**	**UMUX-4**
N	12	12	12	12
Missing	0	0	0	0
Mean	6.33	5.42	5.83	5.08
Median	7.00	6.00	6.00	5.50
Standard deviation	0.985	1.38	0.937	1.62
Minimum	4	3	4	3
Maximum	7	7	7	7

Graphical analysis of constructs

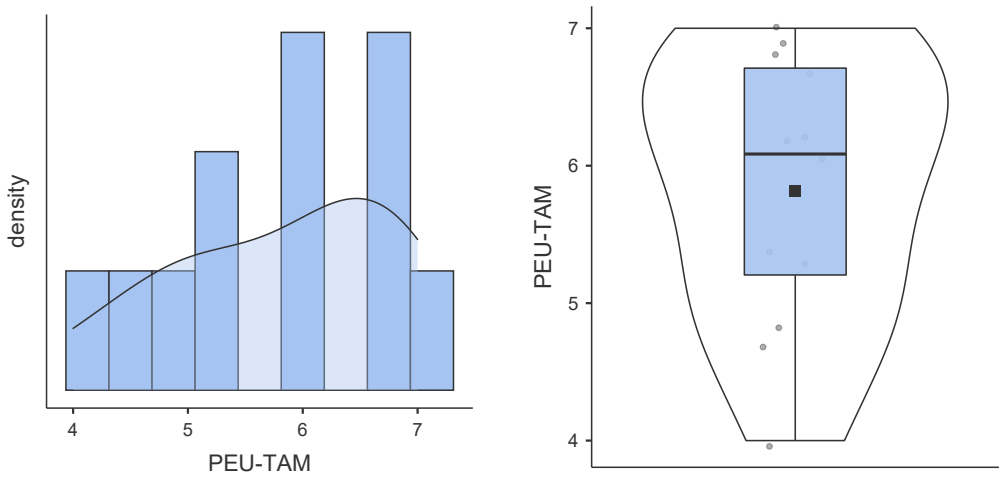
Figures 7.9, 7.10, and 7.11 present the density graphs and boxplots for each of the constructs analyzed, considering the mean scores of each participant. Regarding PU, there is a concentration of ratings between 5 and 7, with a single neutral rating (4). In the boxplot, this neutral rating was highlighted as an outlier because because it differs significantly (lower) from the other data in the set.



(a) Density chart. (b) Boxplot.

Figure 7.9: Perceived Usefulness (PU-TAM) charts.

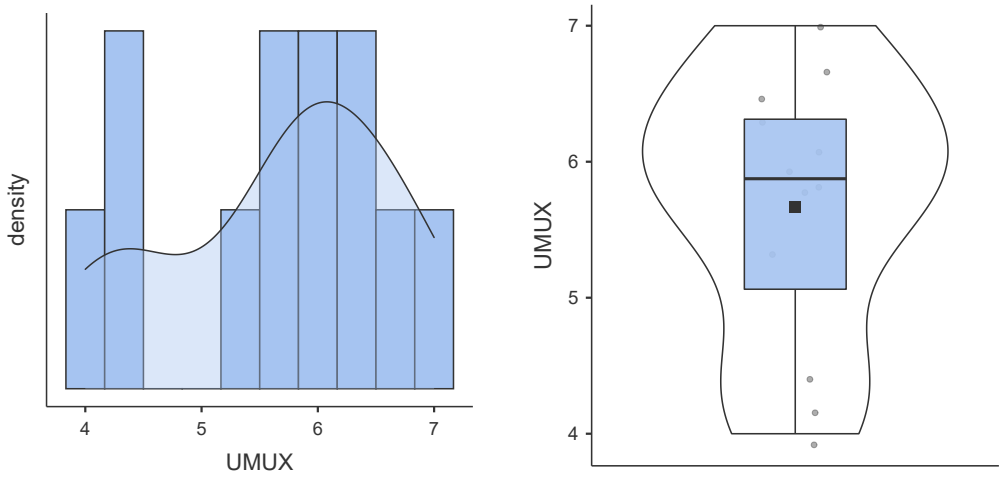
Regarding PEOU, participants' ratings ranged from 4 to 7, increasing the data distribution range, as can be seen in the boxplot.



(a) Density chart. (b) Boxplot.

Figure 7.10: Perceived Ease of Use (PEOU-TAM) charts.

In the UMUX analysis, the behavior was similar to that of PEOU, with the same distribution pattern, despite the slightly lower average.



(a) Density chart. (b) Boxplot.

Figure 7.11: Usability (UMUX) charts.

7.4.3 Correlation analysis between constructs

Table 7.7 presents a summary of the evaluations of each participant in the experimental group, with the average of the evaluations performed for each construct, including the workload, for comparison with the other indicators. The analysis of the data obtained from the application of the TAM and UMUX models indicates a predominantly positive evaluation of the artifact under study.

Table 7.7: Averages of constructs analyzed by participants

ID	Workload	PU-TAM	PEOU-TAM	UMUX
P1	38.00	6.00	6.83	5.75
P2	67.33	5.33	5.33	5.25
P3	71.33	6.50	6.67	6.00
P4	47.66	6.50	6.17	6.75
P5	42.66	5.83	4.67	4.25
P6	51.00	6.67	6.00	5.75
P7	38.33	6.67	6.83	6.50
P8	59.66	6.17	6.17	6.00
P9	40.66	6.17	5.33	4.50
P10	37.00	7.00	7.00	7.00
P11	51.00	3.67	4.00	4.00
P12	41.66	5.83	4.83	6.25
MIN	37.00	3.67	4.00	4.00
MAX	71.33	7.00	7.00	7.00

The means of the constructs associated with TAM were high, indicating high levels of acceptance of the artifact. PU had a mean of 6.03, while PEOU had a mean of 5.82, both on a scale of 1 to 7. These results suggest that participants considered the system useful and easy to interact with. Complementarily, the usability index, measured by UMUX, had an average of 5.67, corroborating the positive perception of use and confirming the convergence between the TAM constructs and the adopted usability metric.

Regarding workload, assessed using NASA-TLX, an average of 48.9 points was found on a scale of 0 to 100, which characterizes a moderate level of cognitive and physical effort required for interaction with the system. However, significant variability was observed among participants: while some reported loads close to 37 points, others indicated values above 70. This dispersion suggests that, although acceptance of the system is high, certain user profiles may have experienced greater effort in performing tasks, pointing to opportunities for improvement in the way the artifact is used. The development of the proposal, in this sense, led to the implementation of support software (the *SafeSecRETS* tool, which had not yet been implemented when the experiment was conducted).

An analysis of the correlation between the constructs used was also performed. The correlation matrix is shown in Figure 7.12.

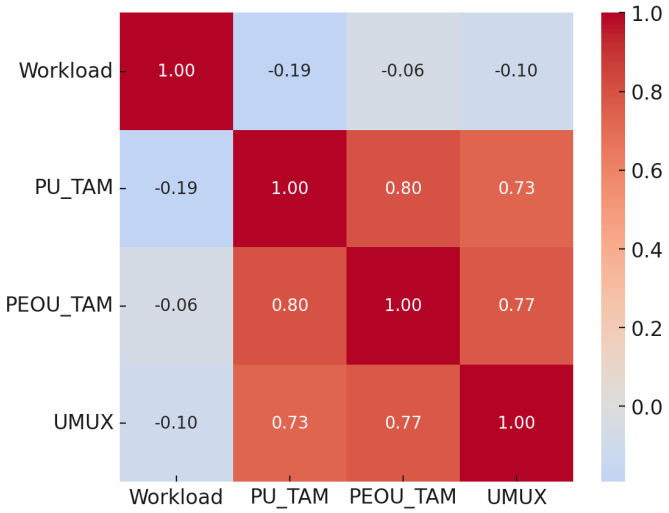


Figure 7.12: Correlation matrix for the constructs evaluated.

Correlation analyses reinforce the theoretical assumptions of the TAM model. A strong association was identified between PU and PEOU (r = 0.80), showing that the perception of ease of use is directly linked to the assessment of the system's usefulness. In addition, both PU and PEOU showed high correlations with the UMUX usability index (r = 0.73 and r = 0.77, respectively), confirming the interdependence between the TAM constructs and the assessment of usability. On the other hand, workload showed negative but weak correlations with the constructs analyzed (ranging from -0.06 to -0.19), indicating that the perceived effort did not significantly compromise the perception of usefulness, ease of use, or usability of the artifact.

7.4.4 Discussion of results for technology acceptance and usability

In summary, the results show that the evaluated artifact (*SafeSecIoT Canvas)* was considered useful, easy to use, and highly usable, confirming its technological acceptance by participants. Although the workload was classified as moderate, this factor did not have a significant impact on the positive perceptions captured by TAM and UMUX. Nevertheless, the individual variability observed in the workload and usability of the participants suggests the need for further investigation into potential improvements in the design of the artifact and interaction processes, with a view to reducing the effort required and improving the user experience.

Regarding the hypotheses, as demonstrated based on the analysis of the results obtained by applying descriptive statistics to the TAM and UMUX data, we can state that:

- "The *SafeSecIoT Canvas* artifact is well accepted in use in supporting the task of developing a project plan for a critical IoT system", **thus refuting hypothesis H02 presented at the beginning of this chapter and validating hypothesis HA02**.

- "The *SafeSecIoT Canvas* artifact has good usability in supporting the task of developing a project plan for a critical IoT system", **thus refuting hypothesis H03 presented at the beginning of this chapter and validating hypothesis HA03**.

Cross-referencing quantitative data and participant feedback

In addition to the closed questions from TAM and UMUX, applied to the experimental group, participants were asked to respond to an open question about positive and negative aspects related to the use of *SafeSecIoT Canvas* for the project planning of a critical IoT system.

A thematic analysis of the responses was performed, and subsequently, the feedback from the participants in the experimental group (open-ended question) was cross-referenced with the quantitative data (generated through the NASA-TLX, TAM, and UMUX questionnaires). This type of analysis allows for a better understanding of the participants' evaluations in each of the defined metrics.

Table 7.8 shows the individual results of this cross-analysis and highlights the positive and negative aspects that were highlighted by each participant.

Table 7.8: Cross-analysis between TAM, UMUX, workload data, and qualitative feedback obtained from open-ended questions to participants.

P	PU-TAM	PEU-TAM	UMUX	Workload	Observations (positive and negative aspects)
P1	High	High	High	Medium	Comprehensive, time demand
P2	High	High	High	High	Complete, laborious
P3	High	High	High	High	Easy to use, for specialists
P4	High	High	High	Medium	Easy to use, many fields
P5	High	Medium	Medium	Medium	Maps security, for complex projects
P6	High	High	High	Medium	Organization, many fields
P7	High	High	High	Medium	Planning, many fields
P8	High	High	High	Medium	Guidance and structure, many fields
P9	High	High	Medium	Medium	Scope definition and organization, —
P10	High	High	High	Medium	Clear and easy to use, —
P11	Medium	Medium	Medium	Medium	Coherence, many fields
P12	High	Medium	High	Medium	Organization and clarity, many fields

Note. Color coding: High, Medium, High (negative).

Overall, PU and PEOU received predominantly "High" evaluations, with only a few "Medium" responses, indicating that the method was generally perceived as useful

and easy to use. Similarly, UMUX results were mostly “High”, although with greater variation, suggesting that usability was positively assessed by most participants.

In contrast, workload emerged as a critical factor: while most participants rated it as “Medium,” two participants reported a “High” workload, and none considered it “Low.” This indicates that, despite its perceived usefulness and usability, the method demanded a considerable cognitive and temporal effort.

The qualitative observations reinforce this interpretation. Positive aspects frequently mentioned include organization, clarity, and comprehensiveness, as well as support for complex projects. Conversely, negative aspects were consistently associated with the large number of fields to be completed and the time required, reflecting a trade-off between richness of information and the effort involved.

In summary, the method was highly valued in terms of PU, PEOU, and usability, but its adoption may be hindered by the perceived workload, particularly for users facing time constraints or less familiar with the domain. The results obtained indicate directions for improvements and future work in relation to the *SafeSecIoT Canvas* artifact. These improvements were addressed after the experiment was conducted, mainly in the development of the *SafeSecRETS* tool.

7.5 Chapter Summary

This chapter was dedicated to the empirical evaluation of the *SafeSecIoT Canvas* artifact, which is the initial and essential element for the safety and security RE process proposed for critical IoT systems. The *SafeSecIoT Canvas* was developed to fill the gap in project planning, assisting in defining the scope, general requirements, and identification of IoT and safety/security concerns from the system’s conception. The evaluation was conducted through a controlled experiment and a survey, applied to undergraduate students in an Internet of Things course.

The results obtained through data collected by the NASA TLX, TAM, and UMUX instruments provided a comprehensive view of the artifact’s effectiveness and acceptance. Below is a summary of the evaluation results:

- Workload (NASA TLX): The controlled experiment compared the experimental group (which used *SafeSecIoT Canvas*) with the control group (which did not use it) in the task of planning a critical IoT system project. The analysis showed that the use of the artifact contributed to a significant reduction in the average workload perceived by the participants in the experimental group, which was 24.74% lower than that of the control group. A positive impact was observed in the perception of lower time demand and frustration, as well as a better performance evaluation

in the task. Thus, the hypothesis that the use of the artifact reduces the workload (HA01_A) was validated.

- Quality and consistency of artifacts: The quantitative-qualitative analysis of the project plans produced by the teams, using techniques such as CFA and Labbé's distance matrix, revealed that the groups that used the *SafeSecIoT Canvas* showed greater semantic similarity and greater consistency with the reference corpus (template) produced by the researchers, compared to the control groups. This finding reinforces that the artifact not only facilitates the task but also increases the effectiveness and quality of the results obtained in project planning.
- Acceptance and usability (TAM and UMUX): The survey conducted with the experimental group indicated a predominantly positive evaluation of the artifact.
  - The PU index reached an average of 6.03 (on a scale of 1 to 7), and the PEOU obtained an average of 5.82, indicating high levels of acceptance.
  - The usability of the artifact, measured with UMUX, was also well evaluated, with an average of 5.67.
  - Both alternative hypotheses related to acceptance (HA02) and usability (HA03) were validated.

Although the *SafeSecIoT Canvas* proved to be a useful, easy-to-use artifact that effectively reduced the average workload, the results also pointed to individual variability in workload (with scores ranging from 37 to 71.33 points on the NASA TLX) and usability. This dispersion suggested that certain user profiles may experience greater effort when interacting with the artifact in its physical form (printed canvas). In line with the qualitative feedback from participants, which indicated aspects to be improved in the interaction process, the empirical evaluation demonstrated the need for support software to refine the artifact's design and interaction processes, with the aim of reducing the effort required and improving the user experience.

Thus, the results of this chapter served as the basis and justification for the development of a support tool that could support and optimize the use of the *SafeSecIoT Canvas* in the subsequent project planning activity. The development of this support tool, called *SafeSecRETS*, was the next proposed step, presented in Chapter 8, aiming to address the identified opportunities for improvement and mitigate the variability in the user experience observed during the experiment.

CHAPTER 8

SafeSecRETS: A Software Tool Supporting Safety and Security RE for Critical IoT Systems

This chapter introduces *SafeSecRETS*, a collaborative web-based tool that supports safety and security RE for critical IoT systems. The tool integrates the *SafeSecIoT Canvas* with the *STPA-SafeSecIoT* method, combining a visual canvas model with an extended version of STPA. This integration offers requirements engineers, domain specialists, and security analysts a shared environment for project planning, elicitation, analysis, and specification. In addition, *SafeSecRETS* structures project information, ensures traceability between planning elements and STPA-based analysis, and supports a systematic RE process - capabilities not sufficiently addressed by existing tools. To demonstrate its applicability, we present its use in the RE of an automated insulin delivery (AID) system.

8.1 Design of the *SafeSecRETS* Tool

SafeSecRETS (acronym for ***Safe****ty and* ***Sec****urity* ***R****equirements* ***E****ngineering for Critical IoT* ***S****ystems*) is a software tool for RE of critical IoT system projects. Built with Bubble and Supabase and grounded in canvas modeling and STPA, it provides a structured workflow for project planning, requirements elicitation, analysis, and specification.

8.1.1 Functional and non-functional requirements

The main functional requirements of the *SafeSecRETS* tool include, but are not limited to: i) user registration and authentication; ii) project management features, including the association of multiple users to the same project to enable collaborative work; iii) canvas-based building blocks to support project planning; iv) interactive forms for structured information collection; v) provision of instructions for completion at each stage; vii) a pipeline for conducting integrated safety and security analysis based on STPA and for specifying these requirements; and vi) visualization of analysis information and its traceability.

The main non-functional requirements guiding the development of the tool are i) usability: the system must provide an intuitive and interactive interface that follows UI/UX best practices, ensuring a smooth user experience for all supported tasks; ii) scalability: the event-driven architecture must handle an increasing number of simultaneous users without performance degradation; iii) security: robust access control and authentication mechanisms must restrict access and modifications to project data to authorized users only; iv) performance: collaborative sessions should support real-time updates through asynchronous WebSocket communication, minimizing latency and database load; v) persistence and reliability: the database must guarantee ACID compliance to ensure secure and consistent operations.

8.1.2 Tool architecture

Figure 8.1 shows the architecture of the *SafeSecRETS* tool, which combines: i) the principles of modularity and separation of responsibilities of layered architecture; and ii) the reactivity and decoupling of event-driven architecture (EDA). The layers of the architecture and their elements are explained below.

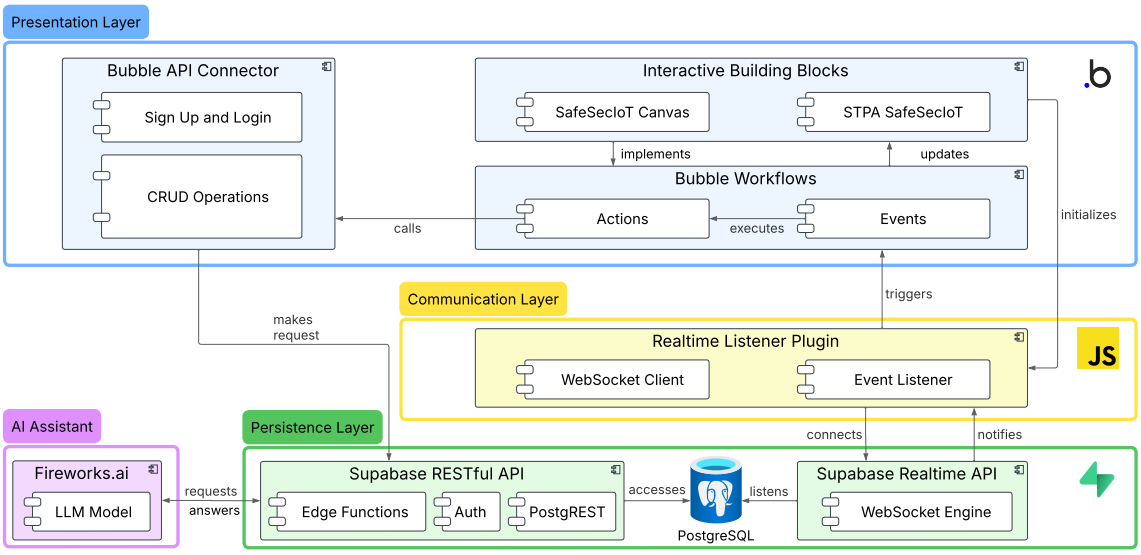


Figure 8.1: Component diagram of the *SafeSecRETS* tool.

- **Presentation Layer**: Manages the tool's graphical interface (*frontend*), allowing users to interact with the building blocks of *SafeSecIoT Canvas* for project planning and STPA-based pipeline. Developed with Bubble, a no-code platform for software development that supports dynamic and interactive interfaces, it defines the workflow logic for processing user events and communicating with the backend. Its main components are i) the *SafeSecIoT Canvas* model: iterative implementation of building blocks for project planning and elicitation of system requirements; ii) the

STPA-SafeSecIoT method: presents the steps for analyzing and specifying safety and security requirements; iii) *Bubble Workflows*: manage navigation, updates, and interactions in the backend; and iv) the *Bubble API Connector*: integrates external APIs such as Supabase for authentication and CRUD operations. User interactions with *SafeSecRETS* trigger workflows, executing actions.

- **Communication Layer (*Realtime Plugin*)**: Built in *JavaScript* as a *Bubble* plugin, it serves as a bridge between the tool's *backend* and *frontend*, allowing *Bubble* to receive real-time database updates. It connects to Supabase Realtime via WebSockets, eliminating constant database queries and improving performance. Its main components are i) the *WebSocket* Client: maintains a persistent connection with *Supabase Realtime*; and ii) the Event Listener: subscribes to specific events (insert, update, delete). When a change occurs in the monitored tables, the plugin triggers an event in Bubble, dynamically updating the user interface with the new information.
- **Persistence Layer**: Handles project data storage and real-time communication (*backend)*. Uses *Supabase*, which integrates a PostgreSQL database with authentication services and REST APIs. Its main components are i) PostgreSQL: relational database; ii) Supabase Realtime: monitors changes in the database and transmits events via WebSockets; and iii) RESTful API: allows authentication and authorization via the Supabase Auth API, ensuring secure access and modification of data. The application developed in Bubble interacts with the database through the Supabase API, while Supabase Realtime notifies the plugin via WebSockets about data updates.
- **AI Assistance Layer**: This is a layer based on AI tools that acts as a provider of a *Large Language Model* (LLM) responsible for automatically generating security restrictions and requirements from specialized information provided by system users. The *Mixtral-8x22B-Instruct* model adopted is widely applied to prompt-based interactive assistance.

8.1.3 Development and technologies

The *SafeSecRETS* tool was developed based on the following technologies: i) *Bubble*, a *no-code* platform for creating web applications with a visual editor for UI design, logic, and database management, along with API integration and *JavaScript* support; ii) Supabase, a backend-as-a-service platform that offers a PostgreSQL database with automatically generated APIs, authentication, storage, and real-time features; iii) Fireworks AI, a platform that offers services for the development and use of generative AI models, focusing on large-scale, low-cost LLM inference; and iv) JavaScript and TypeScript, used to implement scripts and auxiliary plugins (such as Realtime Plugin),

enabling structured communication between Bubble's frontend and backend services. plugins (such as *Realtime Plugin*), enabling structured communication between *Bubble*'s *frontend* and *backend* services and interactions involving LLM-based analysis through *Fireworks AI*.

8.2 Overview of the *SafeSecRETS* Tool

In this section, we present the *SafeSecRETS* tool in greater detail. Figure 8.2 shows the login screen (*index*) for the tool when accessing its web domain[1]. From this screen, a user can create a new account or log in.

Figure 8.2: Index screen of the *SafeSecRETS* tool.

8.2.1 Creating users and projects

Once logged in, the user accesses the tool's home screen, shown in Figure 1. On this screen, they can access previously created projects that have been shared with them, or create a new project. They can also log off from the tool.

Figure 8.4 shows the creation of a new project in the tool. When creating a new project, you must enter a name (title) and other metadata that characterizes the project. This metadata can be edited later, if necessary.

[1] URL for the *SafeSecRETS* tool: https://safesecrets.bubbleapps.io/version-test/

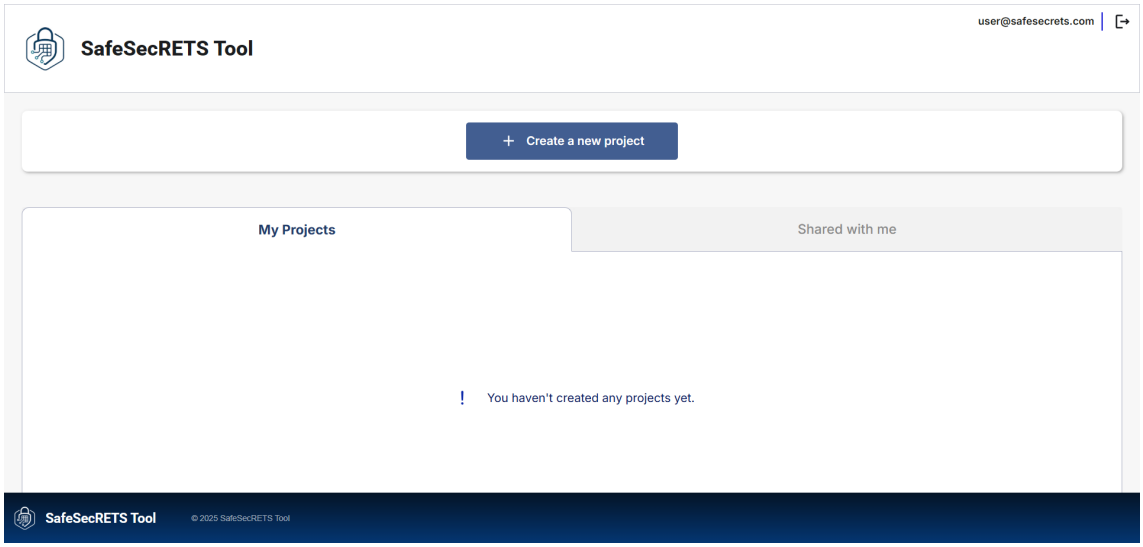


Figure 8.3: Home screen of the *SafeSecRETS* tool.

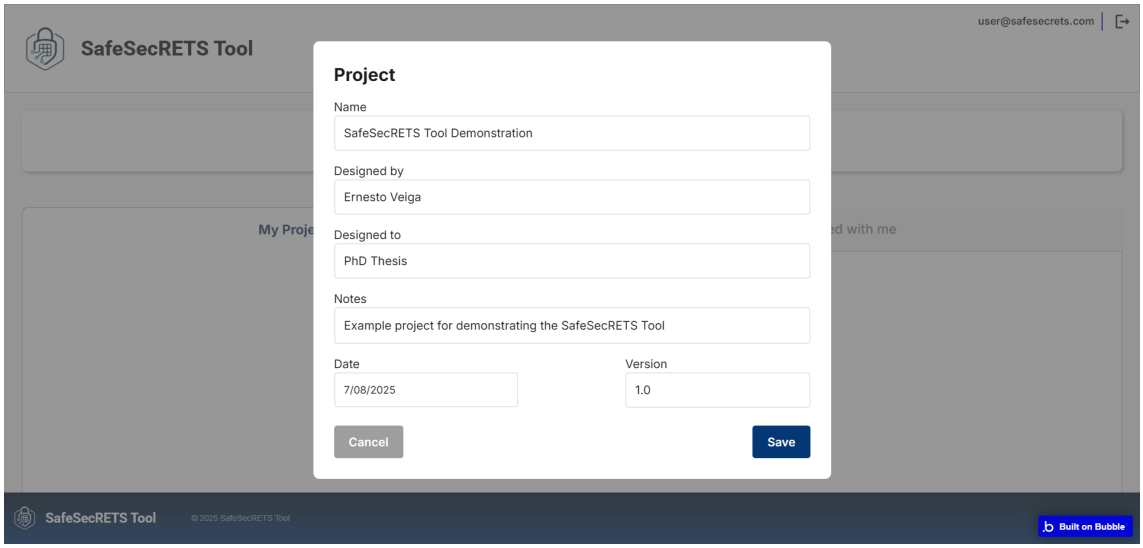


Figure 8.4: Creating a new project in the *SafeSecRETS* tool.

Upon completing the creation of a new project, the user is directed to a page with an overview of the *SafeSecIoT Canvas*, as shown in Figure 8.5.

This page presents a complete view of the *SafeSecIoT Canvas*, with all its building blocks organized according to the fundamental issues they represent (separated by color). In addition, each building block has instructions for filling it out, which will also be available on the fill-in screen for each specific block, as will be shown. This screen will also be used to display the canvas with the information filled in throughout the process.

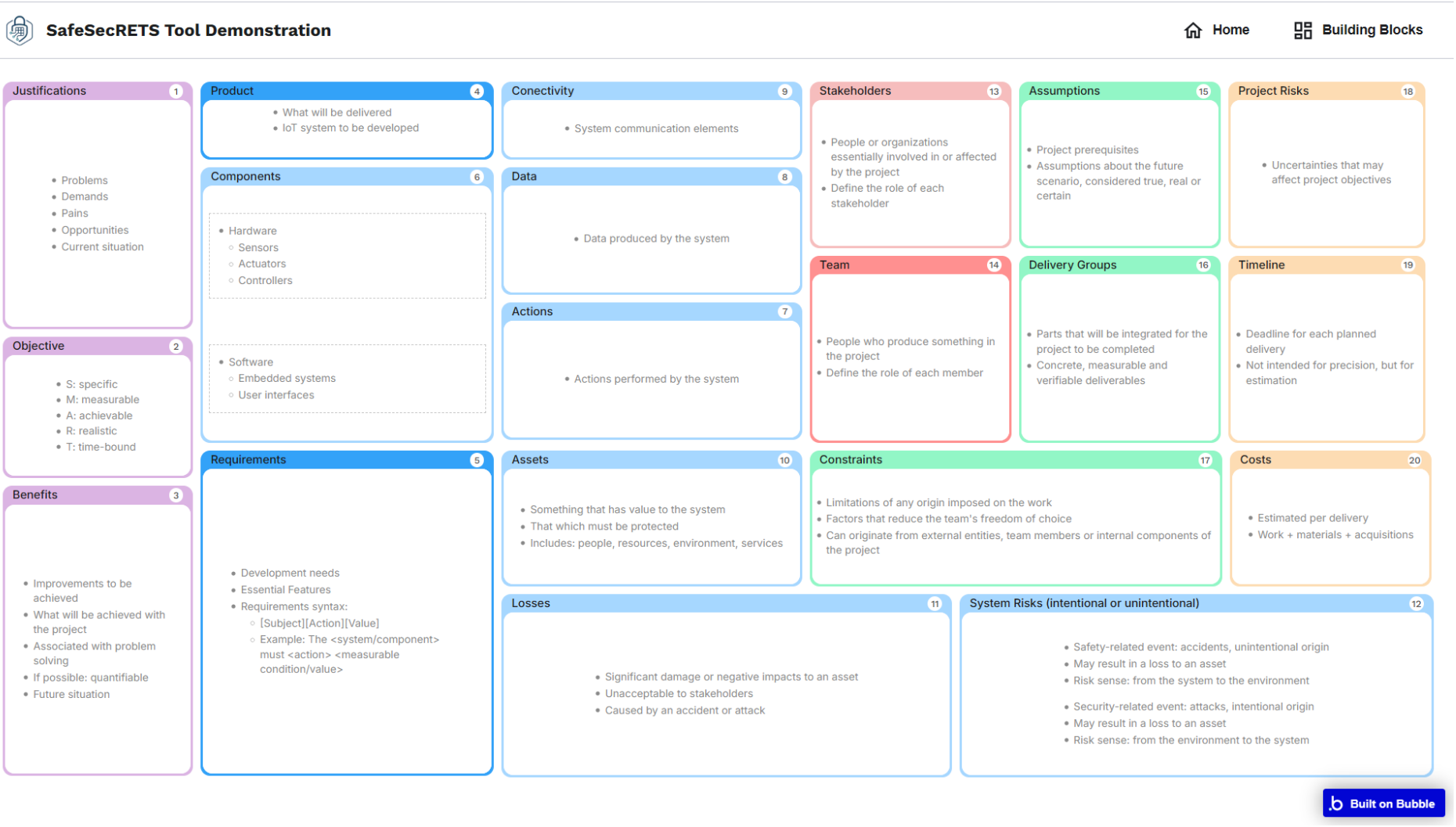


Figure 8.5: Screen showing the *SafeSecIoT Canvas* when creating a new project.

8.2.2 Project planning and requirements elicitation: implementation of the *SafeSecIoT Canvas* model

The *SafeSecRETS* tool implements the *SafeSecIoT Canvas* model, which comprises 20 building blocks for project planning, as shown in Figure 8.5. These blocks are organized into five categories based on fundamental questions (action plan: 5W2H), providing a logical structure for planning. The building blocks within a category are interconnected (based on the conceptual model of the SafeSecIoT Canvas), and may also relate to others outside their category, helping to build relationships between project information, identify inconsistencies, and understand the impacts of changes across the project.

By clicking on any building block in the canvas, the user is directed to the screen of the fundamental question to which it belongs. To facilitate end-user understanding and navigation between the pages of each block group, we have adopted a more specific nomenclature for each group of blocks that make up a fundamental question, working more specifically with the 5W2H constructs.

The first group, *Project Rationale*, presented in Figure 8.6, addresses the "why" of the project (fundamental question: why).

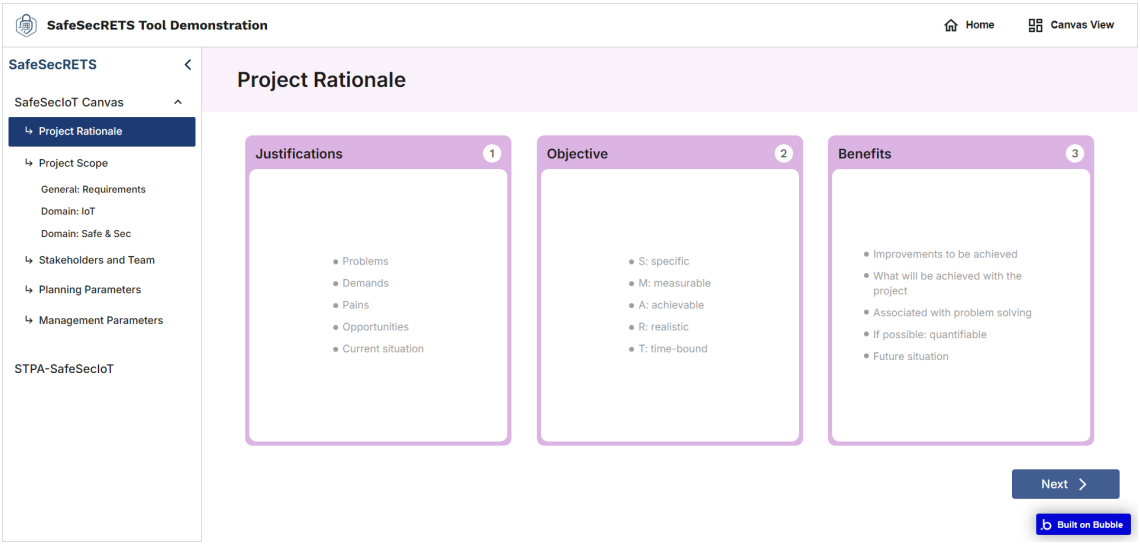


Figure 8.6: Building blocks for project justifications, objectives, and benefits.

The second group, *Project Scope*, defines in more detail "what" should be delivered at the end of the project, the system requirements, and specific domain issues (fundamental question: what), and is divided into three parts:

- Product and system requirements (Figure 8.7);
- Specific elements of the IoT system (Figure 8.8); and
- Specific elements and concerns of critical system safety and security (Figure 8.9).

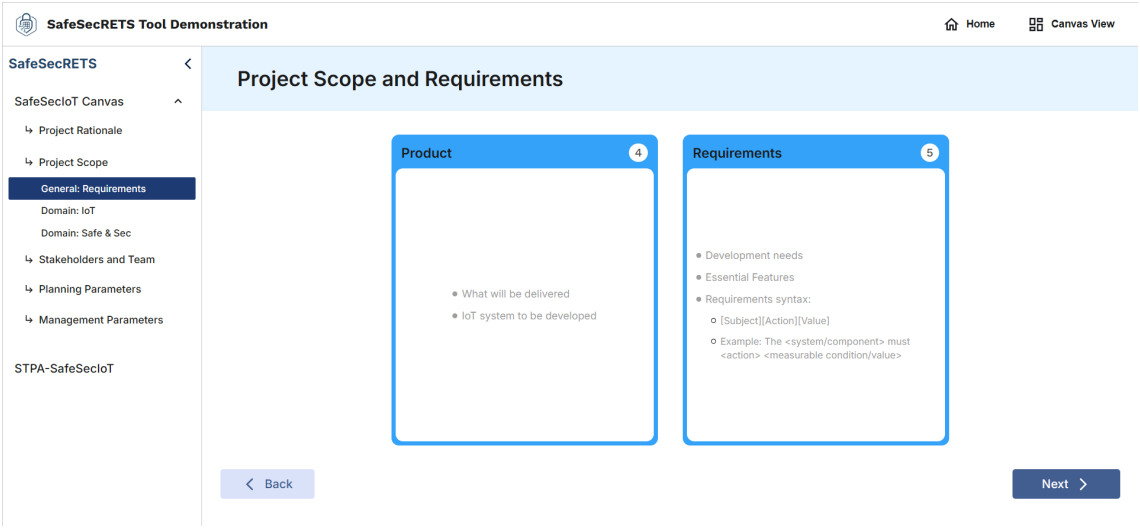


Figure 8.7: Building blocks for defining the product and system requirements.

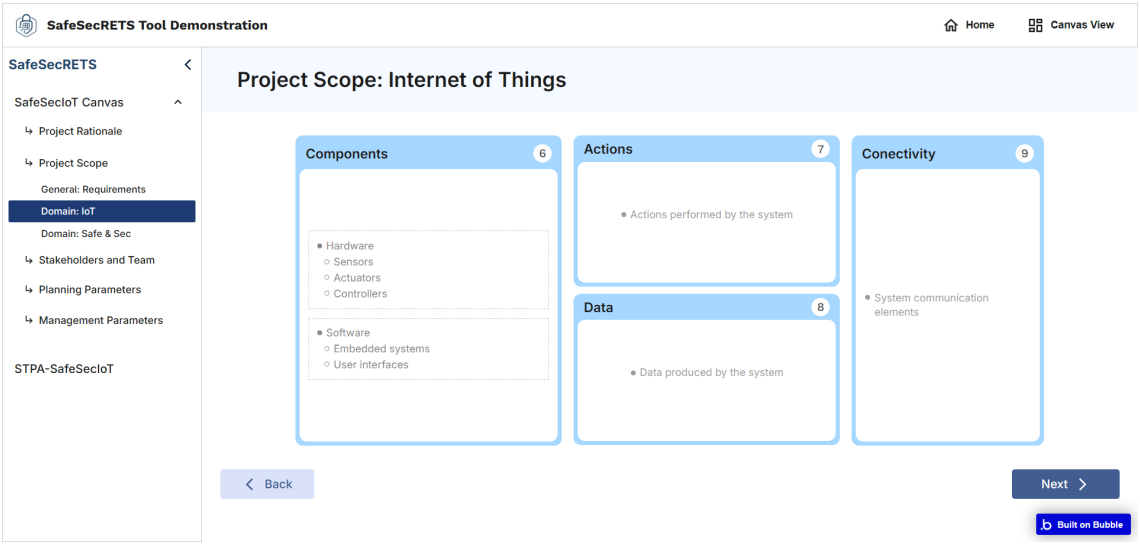


Figure 8.8: Building blocks for defining specific elements of the IoT system.

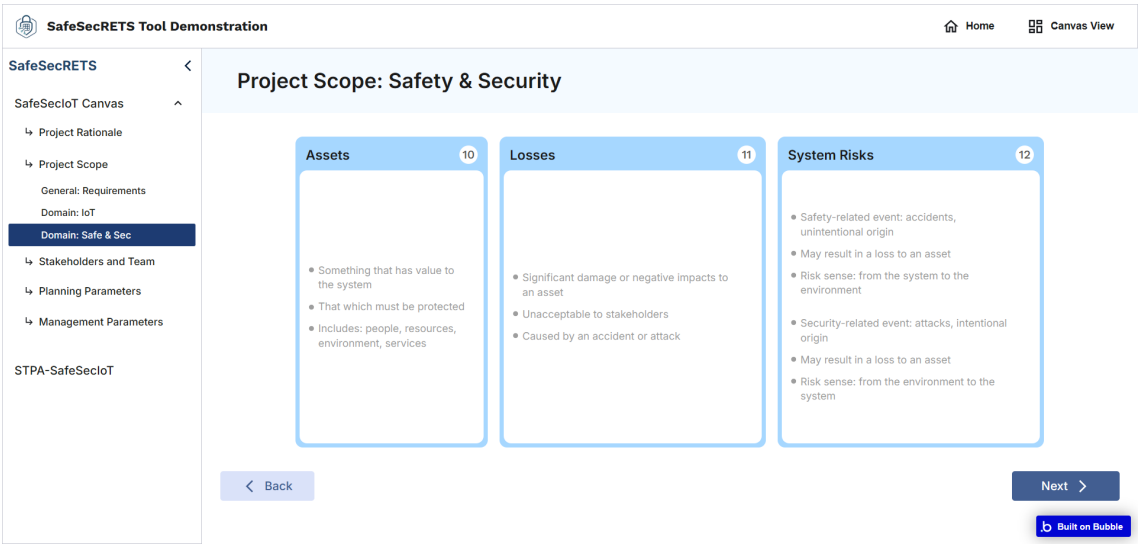


Figure 8.9: Building blocks for defining elements related to system security.

Similarly, the third group, *Stakeholders and Team*, provides building blocks to identify who is involved and their roles in the project. The fourth group, *Planning Parameters*, has blocks for defining "how" the project will be executed, including planned deliverables and constraints. The fifth group, *Management Parameters*, addresses when deliverables should be completed and how much the project will cost, including its potential risks and uncertainties.

As shown in Figure 8.10, each block in a group, when clicked, opens a pop-up window for filling in and editing the necessary information.

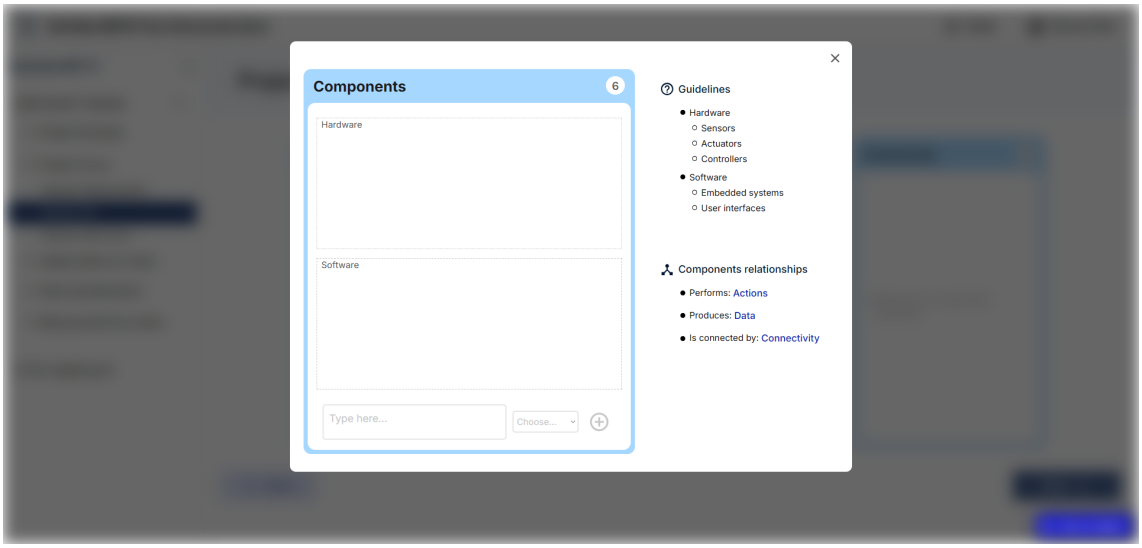


Figure 8.10: Pop-up for editing a building block: instructions and relationships.

In addition to allowing information to be entered and edited, this pop-up window provides specific guidelines for completing the relevant building block and shows its

relationships with other blocks, as defined by the implemented *SafeSecIoT Canvas* model. These relationships are complemented by links that enable direct navigation between related blocks, enhancing the user experience by facilitating interaction with the canvas and access to information.

By structuring project planning around interconnected building blocks that are easy to view and interact with, this implementation of the *SafeSecIoT Canvas* helps users identify, organize, and reason about general project concerns as well as domain-specific ones. The building blocks of the *SafeSecIoT Canvas* include traditional aspects of project planning and management and specialized dimensions, such as specific elements of IoT systems and safety and security concerns, which are essential in critical contexts.

Thus, project planning with the canvas is the starting point for identifying, extracting, and structuring key information that will serve as input for detailed analysis and STPA-based safety and security specification. Thus, the *SafeSecIoT Canvas* bridges the gap between strategic planning and the RE process, allowing the *SafeSecRETS* tool to offer a unified workflow based on methodological rigor and visual guidance.

8.2.3 Analysis and specification of safety and security requirements: implementation of the *STPA-SafeSecIoT* method

To support the analysis and specification of safety and security requirements based on the strategic information defined in the canvas, the *SafeSecRETS* tool implements the *STPA-SafeSecIoT* method in the form of a visual and structured pipeline based on its specific steps and tasks. This method extends the original STPA approach to address safety and security perspectives and requirements in IoT systems, integrating hazard analysis with threat modeling.

This part of the tool aims to guide users through the steps of *STPA-SafeSecIoT* in a systematic and interactive way, performing the tasks related to each one: Step 1) identification of potential losses, hazards and threats, and safety and security constraints at the system level; Step 2) modeling the control structure, including controllers, their responsibilities, control actions, and feedback; and Step 3) identifying unsafe control actions (UCAs) and the corresponding safety and security requirements. Each step is supported by visual elements that represent the central entities of the method and their relationships, mainly traceability.

By strategically linking canvas building blocks to *STPA-SafeSecIoT* elements, the *SafeSecRETS* tool ensures traceability between design decisions and technical analyses. For example, IoT and safety/security concerns defined during planning are used to establish the definition of hazards, threats, controllers, actions, and feedback. Filling instructions and navigation features assist users in performing the steps and tasks of the

process implemented by the tool, meeting the objectives of the analysis until safety and security requirements are obtained, promoting consistency and integrity throughout the analysis.

Since STPA-based analysis depends on project planning information, using some of it as input to pre-fill some steps of the analysis, the tool only enables this part of the process when the necessary information has already been filled in. For this reason, we will present the pipeline (and screen examples) of the implementation of *STPA-SafeSecIoT* in the next section, along with a demonstration of how to use the tool.

8.2.4 Overview about IA Assistant

The AI Assistant module in *SafeSecRETS* automates the generation of part of the analysis information as system-level safety and security constraints (from hazards and threats) and safety and security requirements (from UCAs). Implemented as a Supabase edge function and integrated within the Bubble-based platform, the module accesses the project database to identify UCAs lacking associated requirements and produces formalized specifications in real time.

The module leverages the Fireworks API together with the Mixtral model to transform UCAs into technical requirements. Safety requirements are generated for unsafe control actions, and security requirements for unsecured control actions, following predefined formal templates. The system ensures the generation of precise, traceable requirements that are immediately suitable for use in critical IoT projects. The Figure 8.11 shows an excerpt from the code created for the edge function mentioned above.

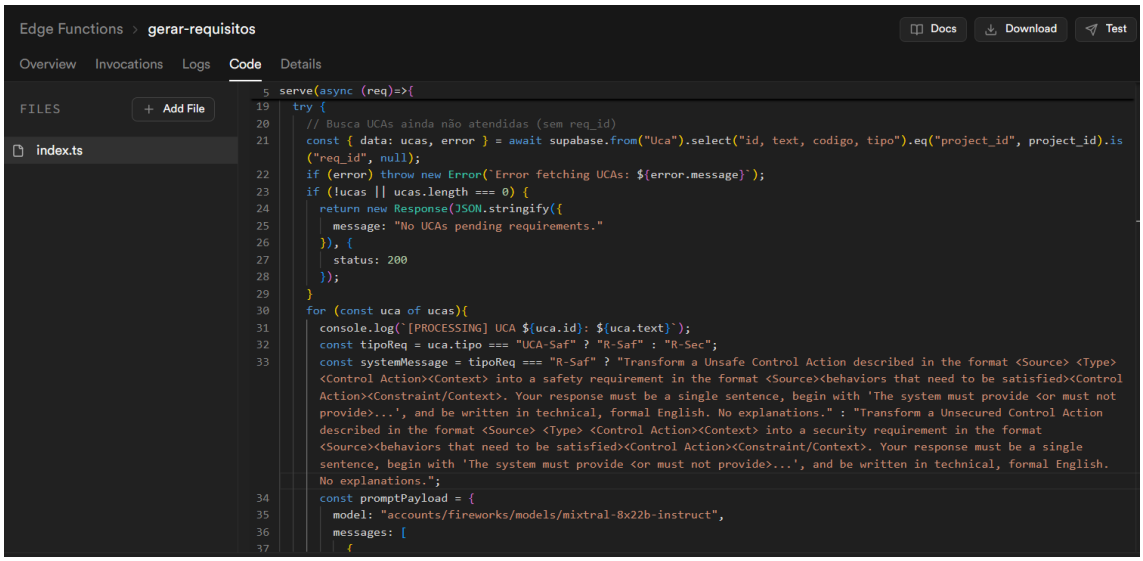


Figure 8.11: Edge function to transform UCAs into technical requirements.

A key feature of the AI Assistant is its flexibility. Although it currently uses the Mixtral model, the module can be adapted to alternative language models with minimal changes. Its design also supports potential extensions, including different project schemas, new requirement types, or integration with other AI-based analysis tools, making it a versatile component for safety and security engineering in critical IoT systems.

8.3 Proof of Concept: demonstration of *SafeSecRETS* tool use in a critical IoT system

To demonstrate the use and potential of the *SafeSecRETS* tool in supporting the safety and security RE process of a critical IoT system, we present a proof of concept that performs all stages of the RE process implemented by the tool, in the application domain of an AID system.

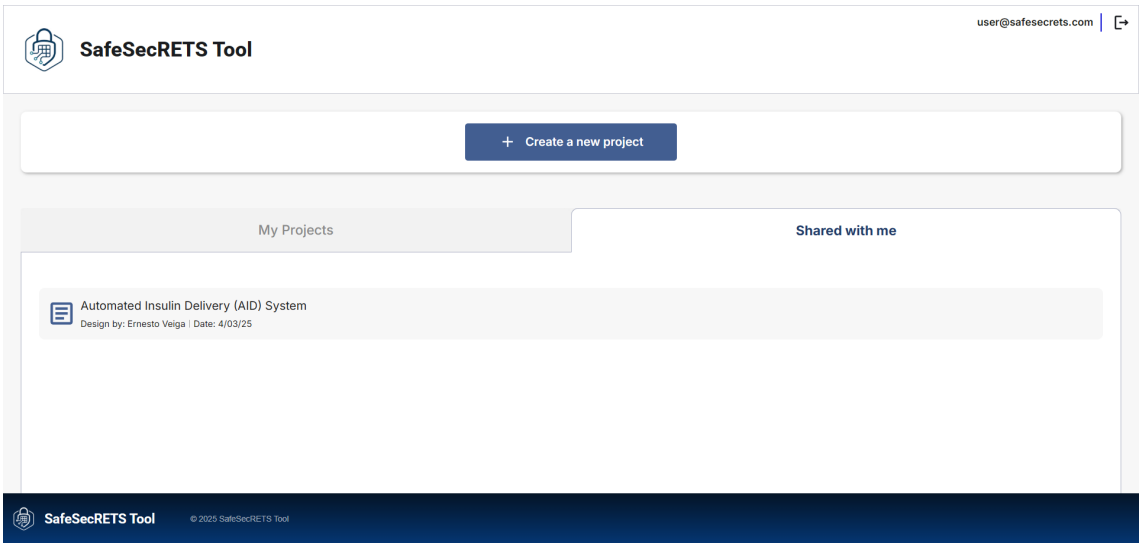


Figure 8.12: Shared project being accessed on the *home* screen.

8.3.1 Project planning support and requirements elicitation

As presented in Subsection 8.2.1, we created a new project in the tool, called *Automated Insulin Delivery (AID) System*. After its creation, this project was shared with a test user of the tool (who has the email address `user@safesecrets.com`). For this reason, the project appears on the *home* screen in the shared projects tab (*Shared with me*), as shown in Figure 8.12.

The AID system project was filled in with the necessary information, including the definition of the system scope (Figure 8.13) and the identification of the essential elements of the IoT system (Figure 8.14) and those related to security (Figure 8.15).

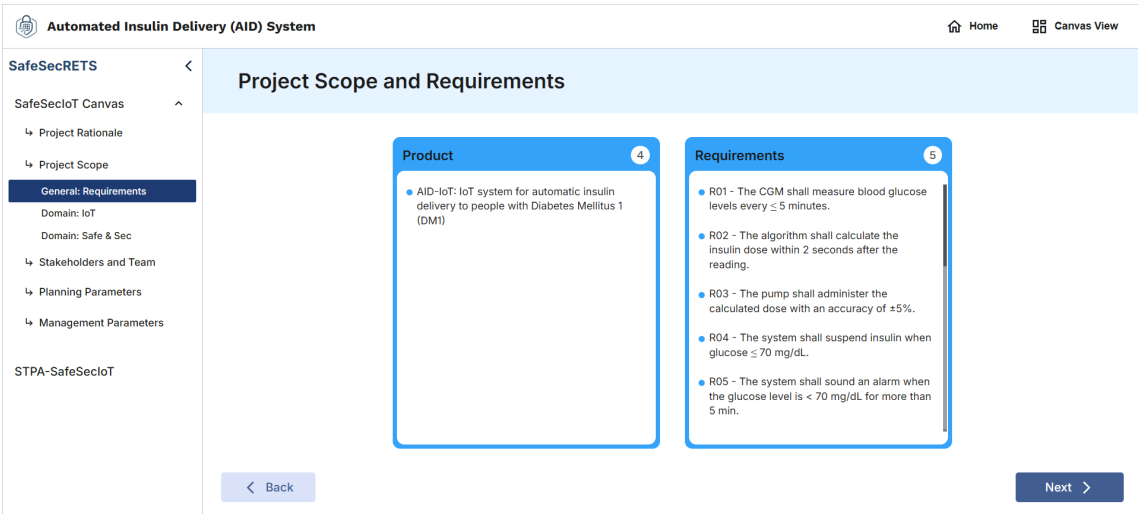


Figure 8.13: Definition of the system scope: description of the product to be developed and elicitation of system requirements.

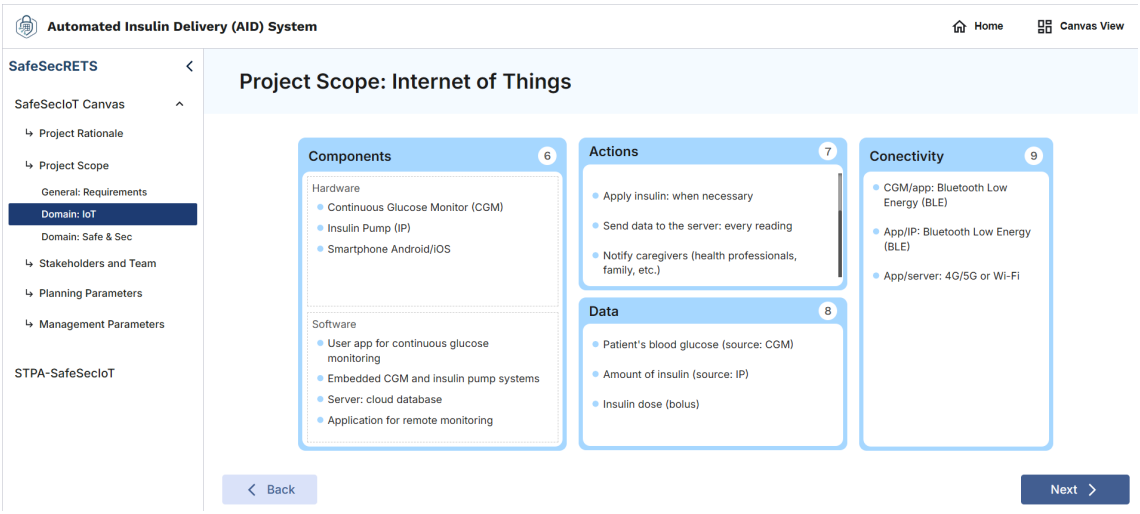


Figure 8.14: Filling in the specific elements of the IoT system domain.

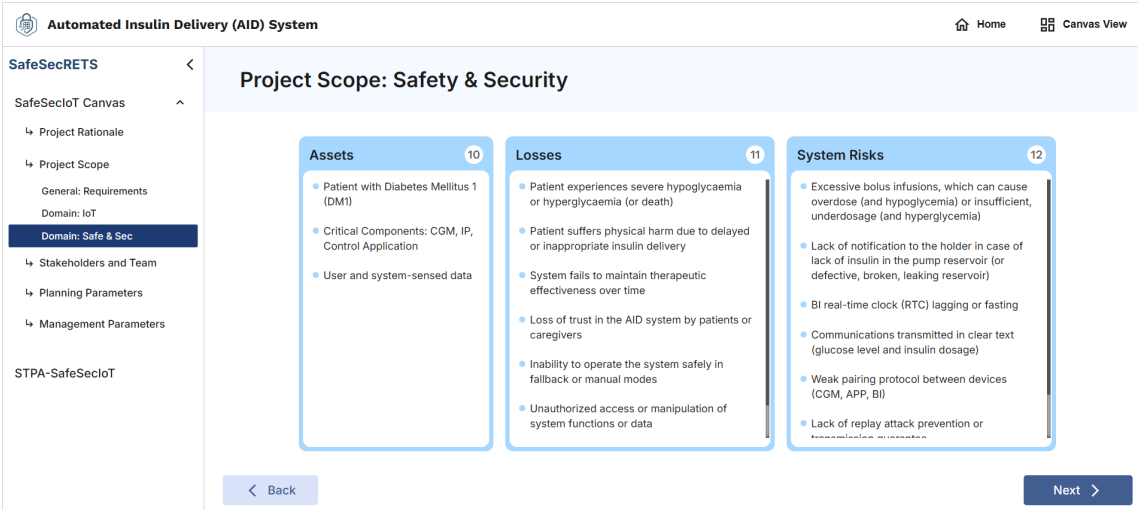


Figure 8.15: Filling in specific elements of the safety and security domain.

Figure 8.16 shows the pop-up for filling in the "Losses" building block. This information, for example, will be used as input in the safety and security analysis process that will be carried out later. Possible losses are high-level information that can be identified during the design of a critical system and will serve as the basis for a hazard and threat analysis and for defining safety and security constraints for the system.

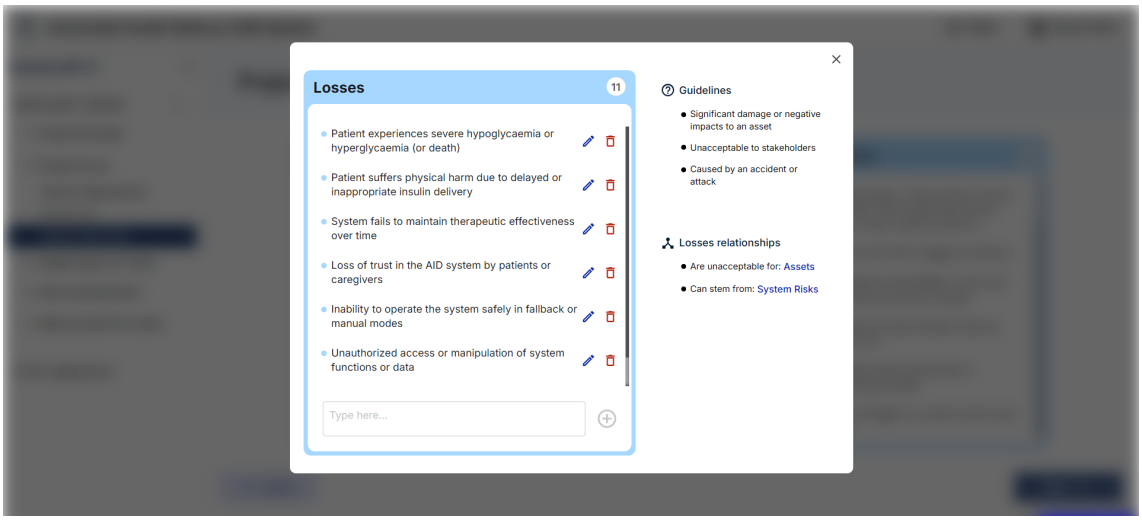


Figure 8.16: Editing a building block.

The canvas with all the information needed to start the STPA-based analysis is shown in Figure 8.17. The building blocks that have already been filled in contain the project planning information and description of the AID system. The blocks that have not yet been filled in contain the guidelines/instructions for filling them in.

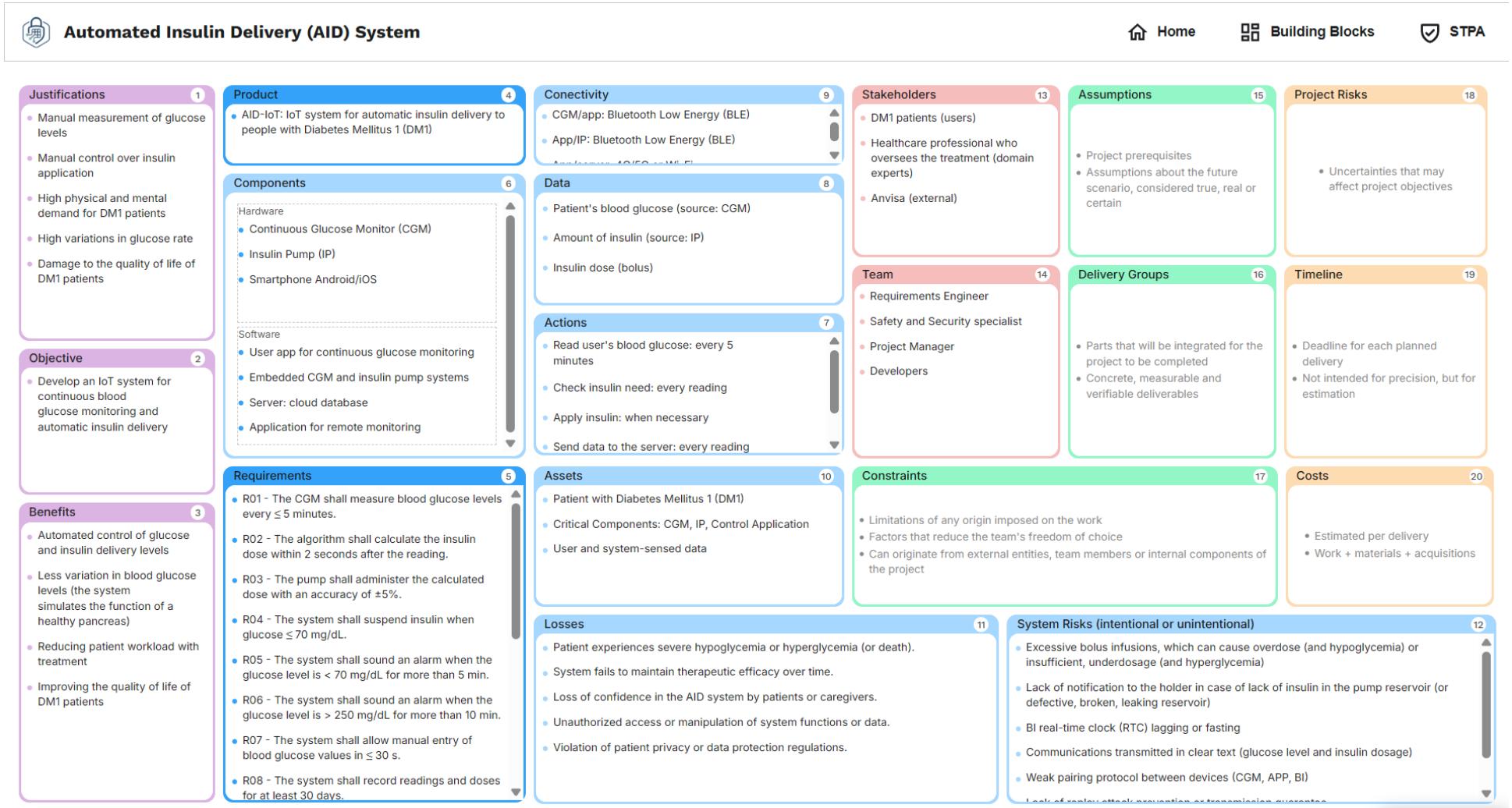


Figure 8.17: Screen showing the *SafeSecIoT Canvas* for the design of an automatic insulin delivery system.

Once all the building blocks that comprise the **Project Rationale** and **Project Scope** have been completed, STPA-based analysis is enabled. The *SafeSecRETS* tool implements the *STPA-SafeSecIoT* method for safety and security analysis and specification, whose steps and tasks can be accessed from the canvas page (STPA button) or from the tool's side menu ("*STPA-SafeSecIoT*" option), from any of the building block groups.

8.3.2 Support for the analysis and specification of safety and security requirements

This section presents the implementation of the *STPA-SafeSecIoT* method in the context of the *SafeSecRETS* tool. This part of the tool aims to support the process of safety and security analysis and specification of these requirements.

Step 1 - Defining the Purpose of the Analysis

The implementation of the first stage of *STPA-SafeSecIoT* uses information obtained through the canvas as inputs for the analysis process. The first task in this stage consists of associating possible losses with system assets already identified in the specific building blocks of the security domain. Figure 8.18 presents the list of losses identified in the context of system planning, already associated with the respective assets.

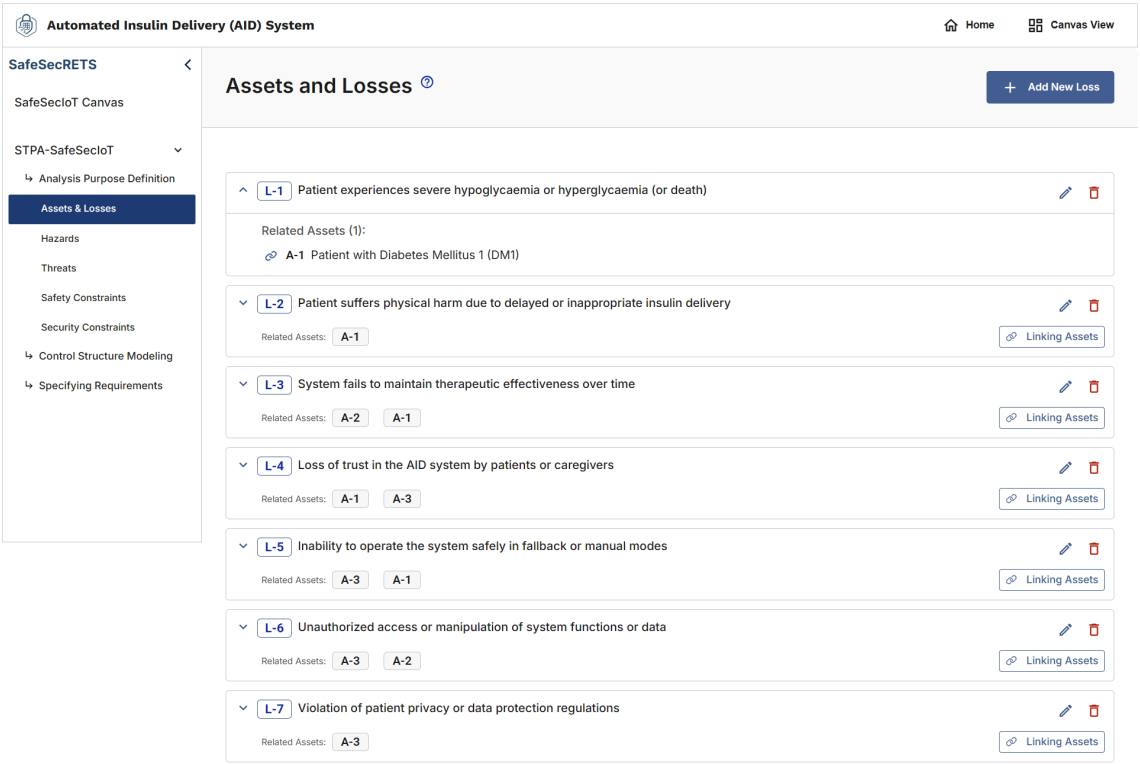


Figure 8.18: First stage of *STPA-SafeSecIoT*: presentation of assets and losses.

Also in Figure 8.18, we demonstrate how each loss can be expanded for a detailed view of the associated asset(s), as exemplified in loss L-1, which is associated with asset A-1. Assets can be linked to losses using the *"Linking assets"* button, which opens a *pop-up* with all assets, allowing the user to define those related to the loss being addressed, as shown in Figure 8.19.

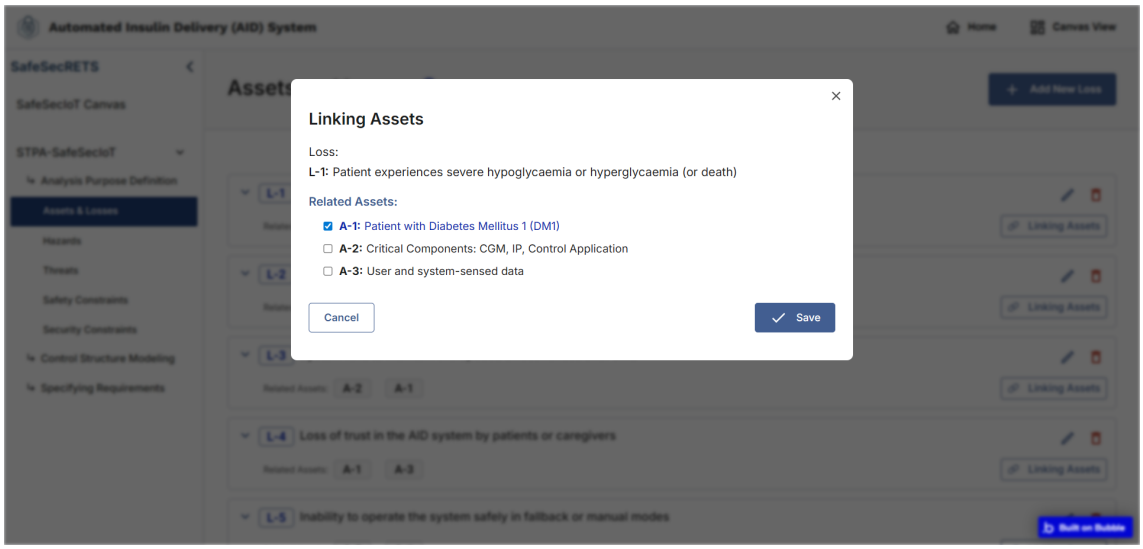


Figure 8.19: Asset association with losses.

If a loss has not been identified when filling in the canvas building blocks, the analyst can add this loss directly on the *"Assets and Losses'"* page, using the *"Add New Loss"* option, as shown in Figure 8.20, or even on the canvas itself. Even if added outside the canvas, this new loss will also be displayed in its respective building block.

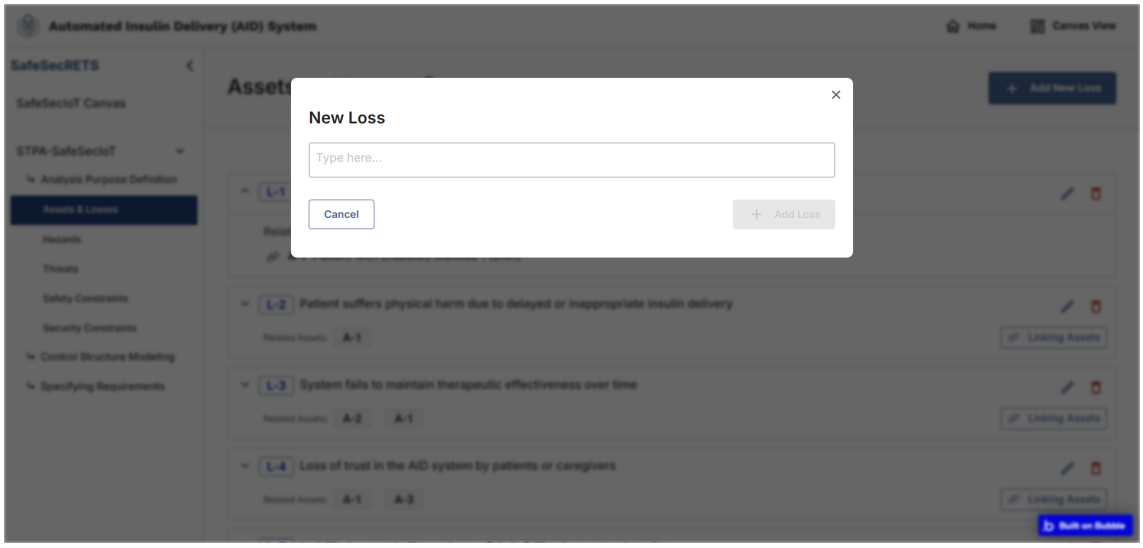


Figure 8.20: Creating a new loss outside the canvas building block.

After identifying potential losses to the system and associating them with the respective assets, the tool offers a pipeline for safety and security analysis based on the information already identified. The first step in this detailed analysis process is to identify hazards (related to safety) and threats to the system (related to security). These hazards and threats can, in the worst case, lead to losses. Therefore, the tool requires the analyst to establish this traceability relationship when creating a new hazard or threat, as shown in Figure 8.21.

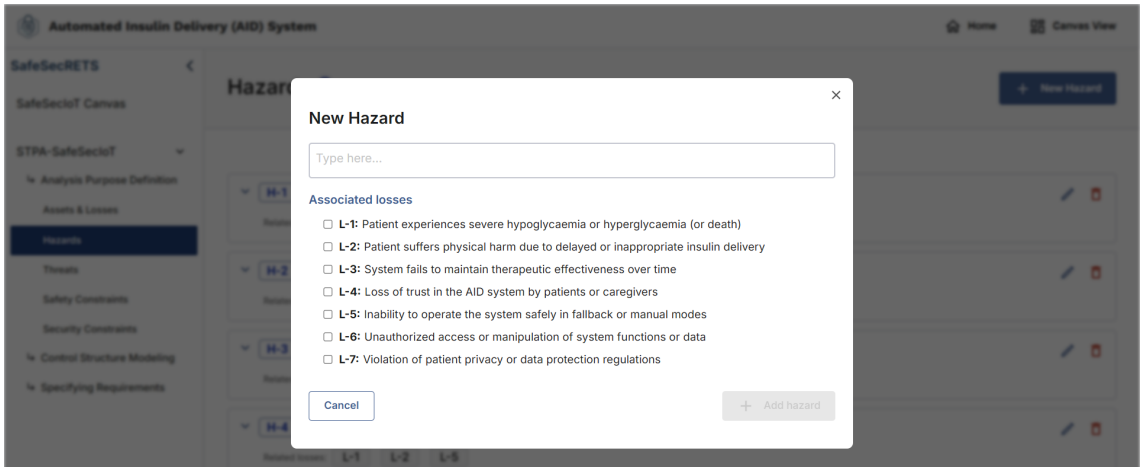


Figure 8.21: Creation of a new hazard and its association with possible losses.

Figure 8.22 shows the screen with the hazards identified for the AID system.

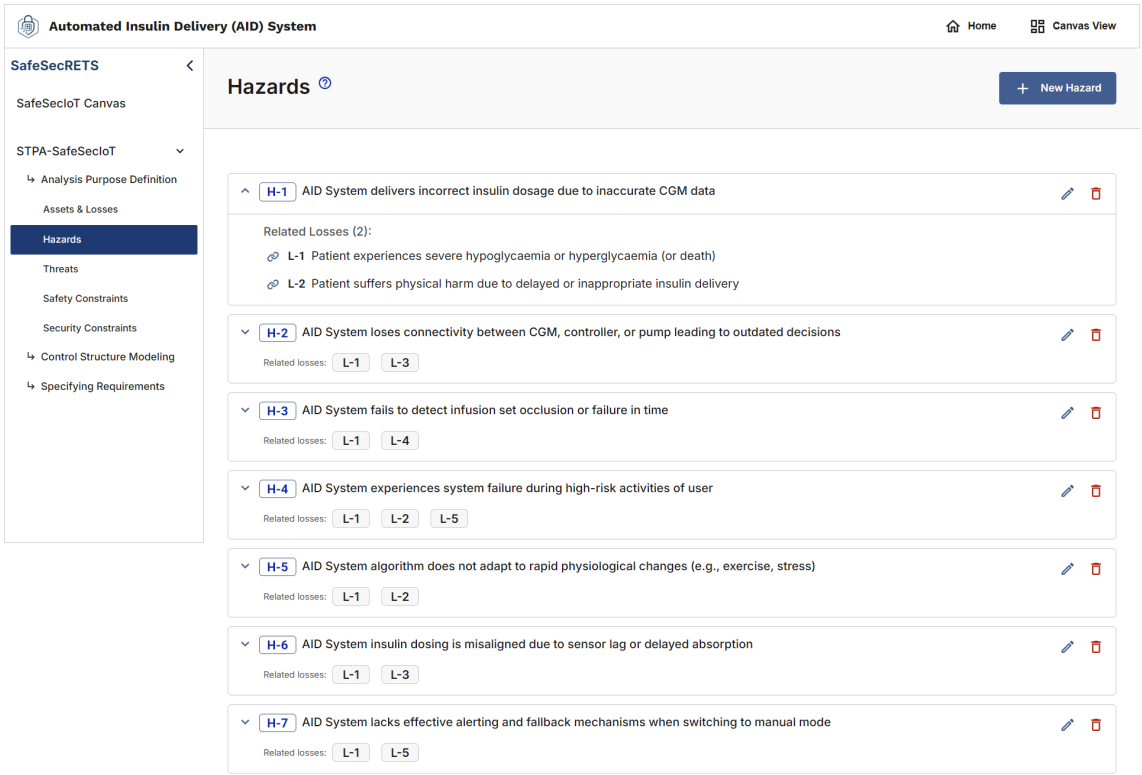


Figure 8.22: List of hazards identified for the AID system, associated with losses.

The same type of interface used for creating and listing hazards is also used for defining threats. Threats represent security risks to the system and should also be related to the possible losses that may occur in a worst-case scenario if they are not properly addressed.

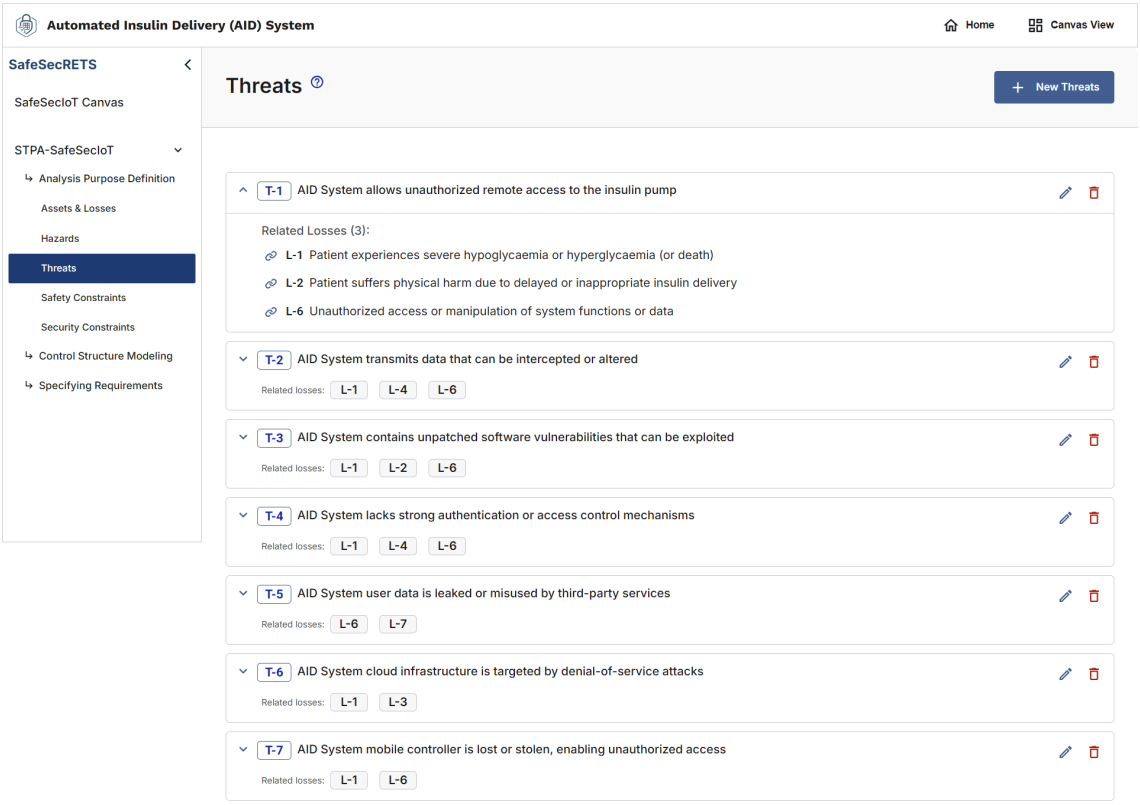


Figure 8.23: List of threats identified for the AID system, associated with losses.

After identifying the potential hazards and threats to the system, the final task in the first stage is to derive safety and security constraints, still at the system level. These constraints are necessary measures to mitigate the potential hazards and threats identified, preventing them from causing a loss to the system.

Since the constraints are derived from the hazards and threats already identified by the analyst(s), we have implemented an AI assistant to assist in the process of automatically generating constraints. This way, analysts can create new safety or security constraints manually, or generate these constraints automatically using the AI assistant of the SafeSecRETS tool.

Figures 8.24 and 8.23 show the screens with system-level safety and security constraints, respectively. In addition to the constraints already generated, there are two buttons in the upper left corner, one for automatically generating constraints using AI, and the second for manually adding a constraint, according to the screen the analyst is on. Each safety or security constraint must be associated with a hazard or threat, respectively, maintaining traceability.

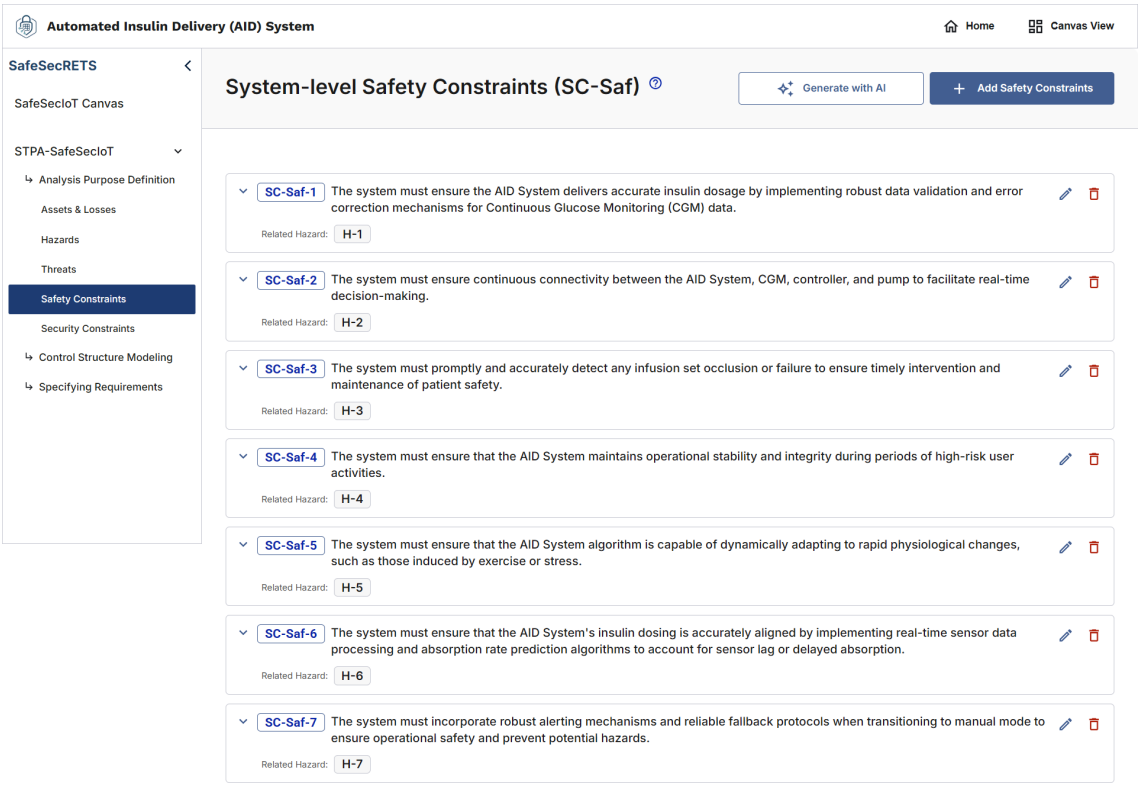


Figure 8.24: Safety restrictions of the AID system, generated by the AI assistant.

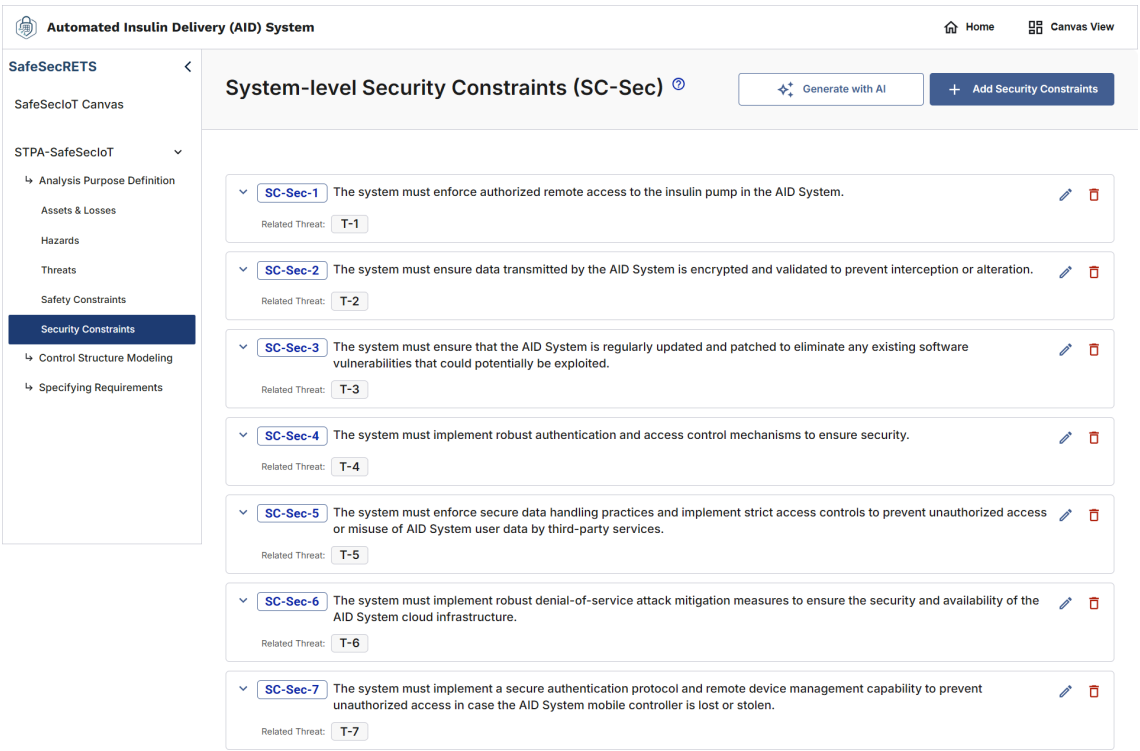


Figure 8.25: AID system security restrictions, generated by the AI assistant.

Step 2 - Modeling the Control Structure

The implementation of the second stage of the *STPA-SafeSecIoT* method, in the context of the *SafeSecRETS* tool, also uses information defined in the canvas as inputs that assist in defining the elements of the system control structure. The first task performed by the analyst in this stage is to define the elements of the control structure, which are based on the IoT system components identified through the *"Components"* building block when filling out the canvas.

Figure 8.26 shows the screen that lists all the controllers created during the definition of the system control structure. It can be seen in the image that each controller, in addition to its code and name, has other related information: i) the associated responsibilities, and ii) the control actions received and issued by the controller, which will be explained below. In addition, to the right of each controller entry, there are two buttons that allow you to add i) a new responsibility or ii) a new control action to the controller being defined.

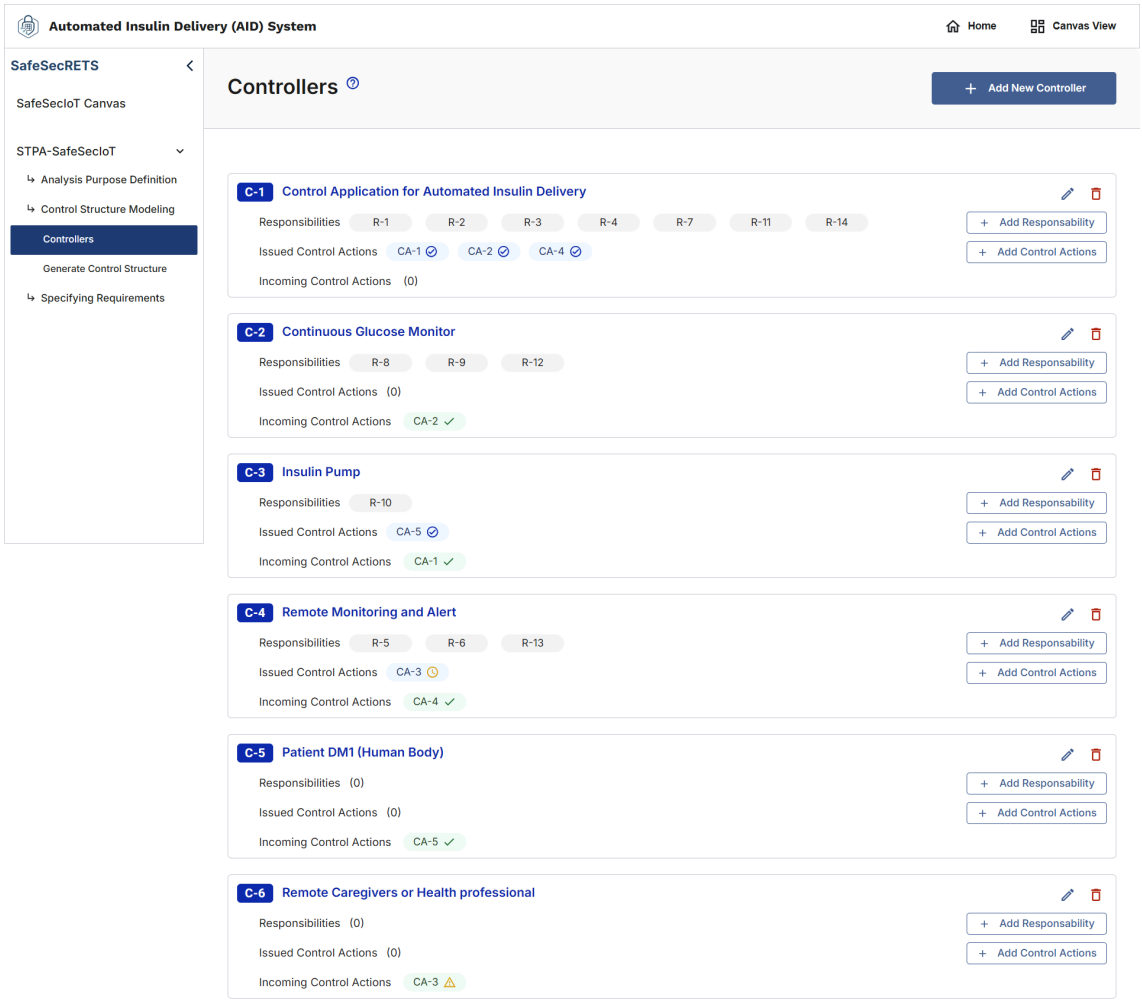


Figure 8.26: Screen for creating and displaying the controllers (and controlled elements) of the analyzed system.

When creating a new controller, it must be associated with a system component defined during project planning as part of the critical IoT system scope. For this reason, a list of all components is presented to the analyst via a pop-up window that the system opens when creating a new controller, as shown in Figure 8.27.

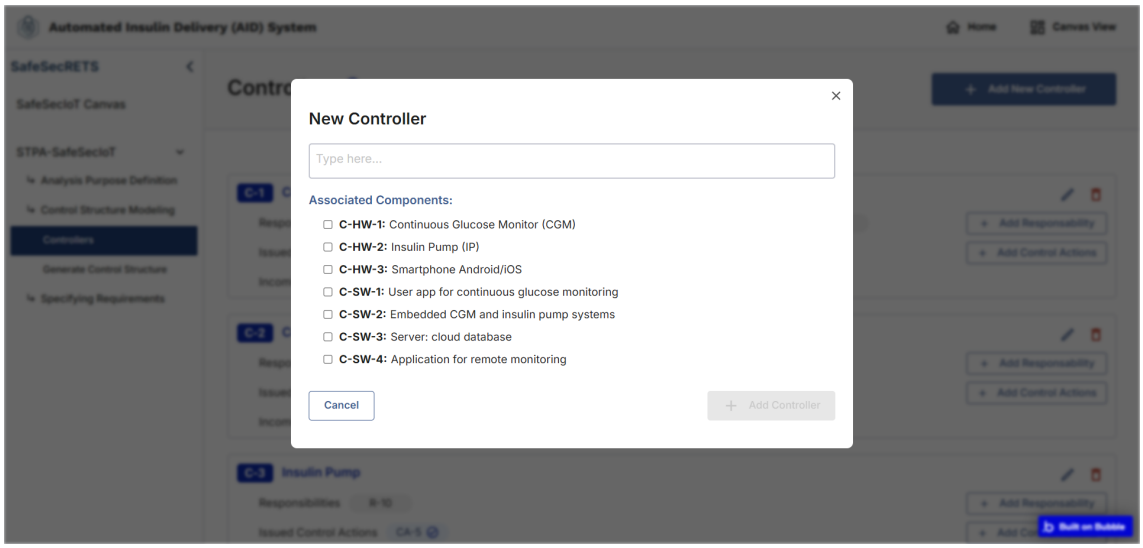


Figure 8.27: Pop-up for creating a new controller.

Once a new controller has been created, the analyst must define the responsibilities related to it within the context of the system. The responsibilities of a controller are associated with a previously defined system-level safety or security constraint. For this reason, these constraints are presented to the analyst, who must associate them as necessary with the responsibility being defined, as shown in Figure 8.28.

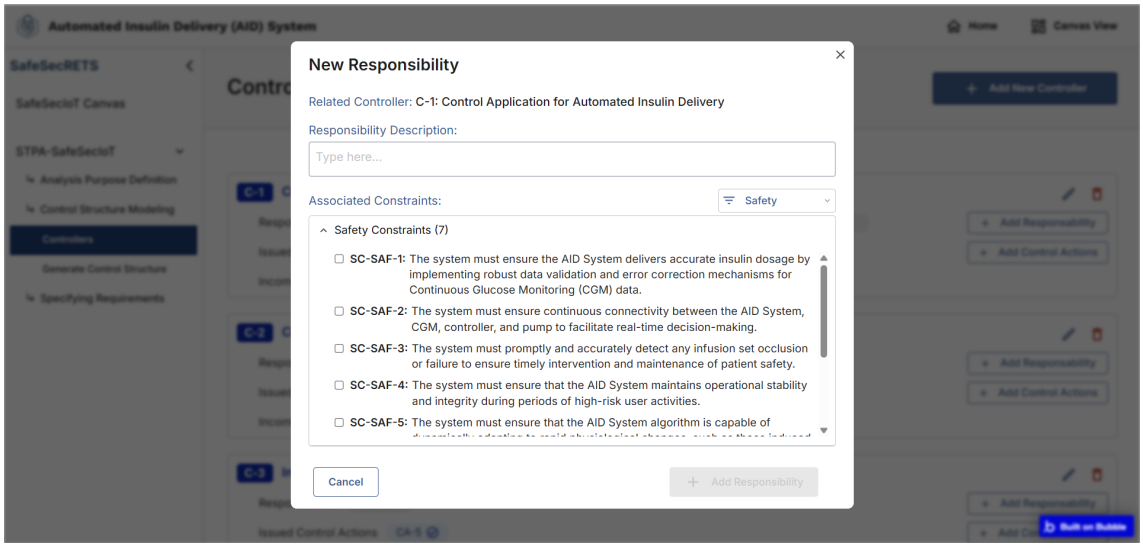


Figure 8.28: Definition of a responsibility for a controller.

Similar to the definition of responsibilities, the tool also offers a feature for defining system control actions, which will be derived from and associated with an action defined in the canvas (building block “Actions”). Once this control action is defined from a controller, it will be associated as the source of the action. Therefore, the tool offers a selection list for defining the controller or controlled element that will be the target of this control action. In addition, it is necessary to associate the control action being created with one or more responsibilities that will be fulfilled by it.

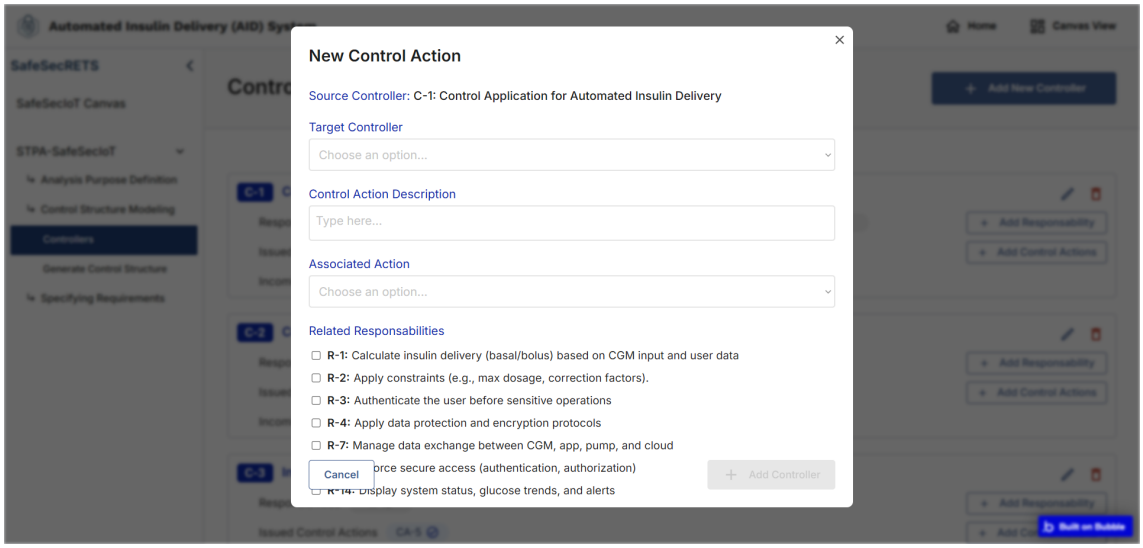


Figure 8.29: Definition of a control action for a controller.

After creating a new controller, it can be accessed on an information display screen, as shown in Figures 8.30 and 8.31. This screen has a set of tabs that, when accessed, allow for a detailed display of all information related to the controller: i) related components; ii) responsibilities associated with this controller; iii) the control actions issued by it (Figure 8.30); iv) the control actions received; and v) the feedback sent to other controllers (Figure 8.31).

In the example shown in Figure 8.30, we can see that the controller in question (*Control Application*) performs the control actions: i) to release insulin delivery (*Release insulin delivery*) and ii) to measure glucose level (*Measure glucose level*). These control actions target (or are directed at) the insulin pump (controller C-3) and the continuous glucose monitor (controller C-2), respectively. The figure also shows details of the other elements related to the control actions: related IoT system action, responsibilities, and feedback (including the target control action and the data sent).

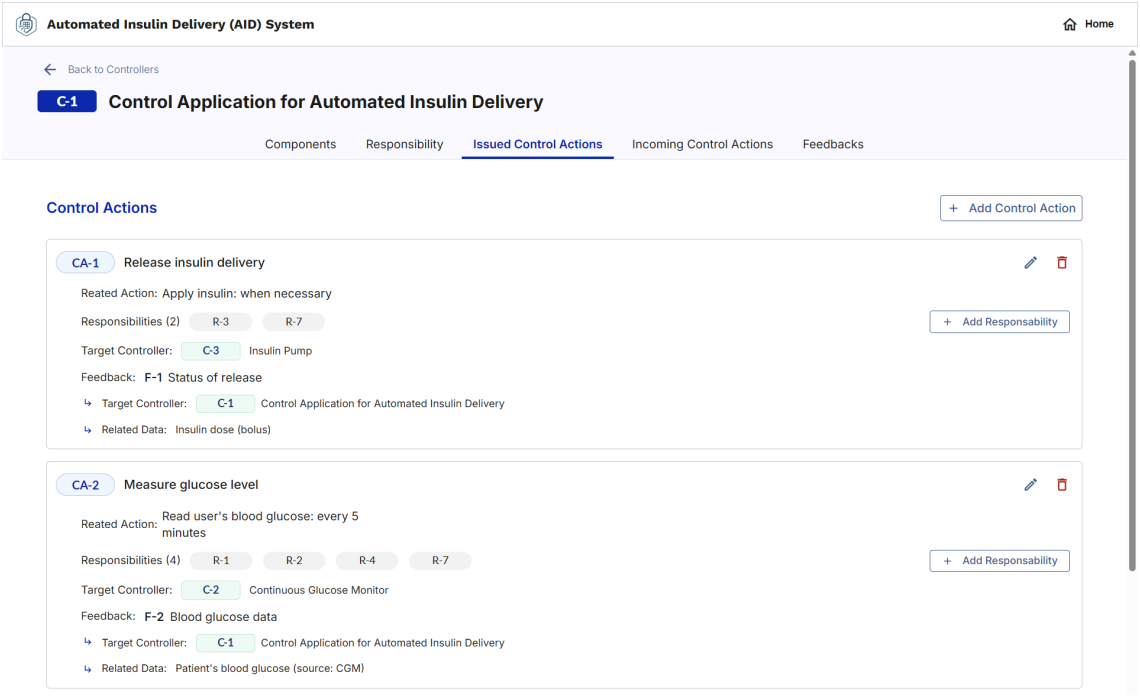


Figure 8.30: Details of a controller with its associated information: highlight the control actions issued.

In the excerpt shown in Figure 8.31, two types of feedback are presented: i) blood glucose data, originating from the continuous glucose monitor; and ii) delivery status, originating from the insulin pump.

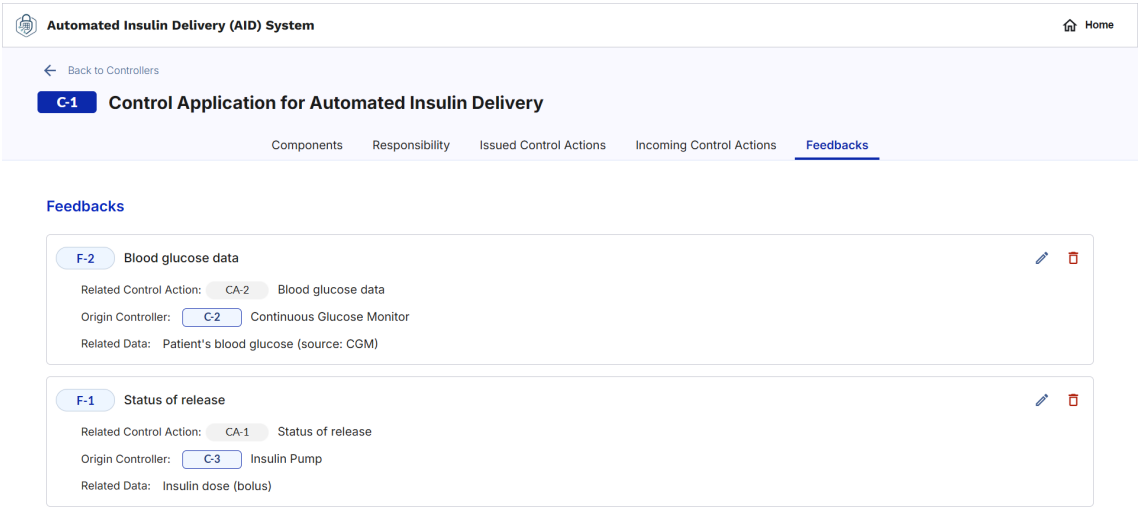


Figure 8.31: Details of a controller with its associated information: highlight the feedback sent.

Based on the definition of the control structure elements, as presented so far, the control structure can be automatically generated by the *SafeSecRETS* tool. Figure 8.32 shows the control structure generated for the AID system.

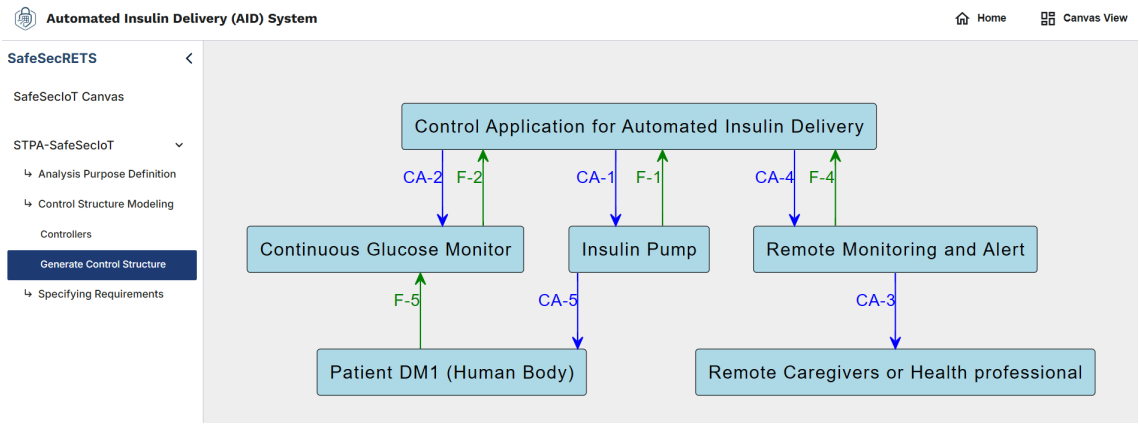


Figure 8.32: Control structure of the AID system automatically generated by the *SafeSecRETS* tool.

Step 3 - Definition of Unsafe/Unsecured Control Actions and Safety and Security Requirements

After defining the system control structure, with its detailed information, the third step of *STPA-SafeSecIoT* involves defining unsafe control actions and specifying the corresponding safety and security requirements. In accordance with the STPA method,

When clicking on the *"Create a New UCA"* button, the user is directed to the screen shown in Figure 8.33. This screen displays a list of the system's control actions, with their associated responsibilities (which can be detailed as needed), allowing a new UCA to be defined to indicate a possible vulnerability in the system related to a specific control action.

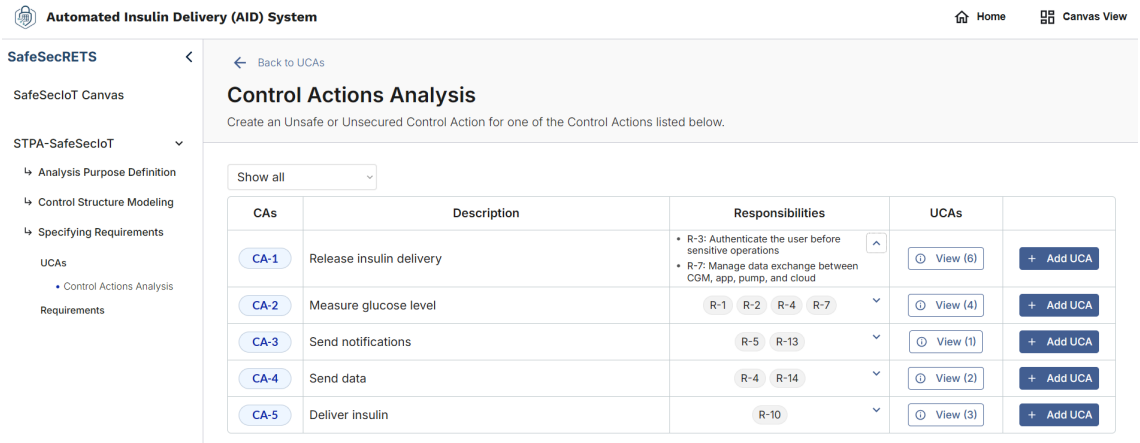


Figure 8.33: Screen for analyzing and defining UCAs.

Clicking the *"Add UCA"* button opens a *pop-up* for specifying the unsafe control action and defining its type, i.e., whether it is related to safety or security. When choosing the type of UCA to be specified, a list of hazards (if it is an *Unsafe Control Action*) or

threats (if it is an *Unsecured Control Action*) is displayed so that they can be associated with the UCA being specified. Figure 8.34 shows the pop-up for defining a UCA.

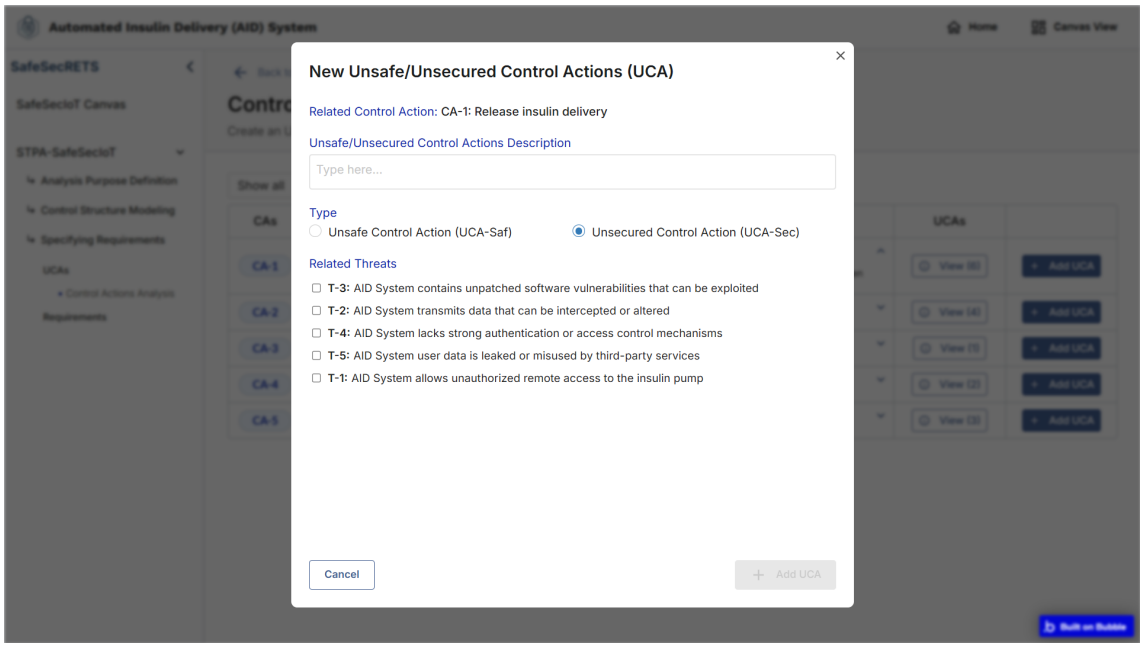


Figure 8.34: Screen for specifying an unsafe/unsecured control action (UCAs).

Figure 8.35 shows the list of UCAs in the system related to safety. All created UCAs are presented in a list of items with the option to expand the information, in the same pattern as the screens for hazards, threats, and safety and security restrictions. Similarly, Figure 8.36 shows the list of UCAs related to the security perspective.

Expanding an item in the list allows you to view the hazards or threats related to each UCA. In addition, this list has a filter to allow you to view UCAs related to safety and security together or separately. Each item in the UCA list also provides information on the related control action (code and description), as well as the number and codes of related hazards or threats. It also has an option to edit the UCA specification and related information or even delete it, if necessary.

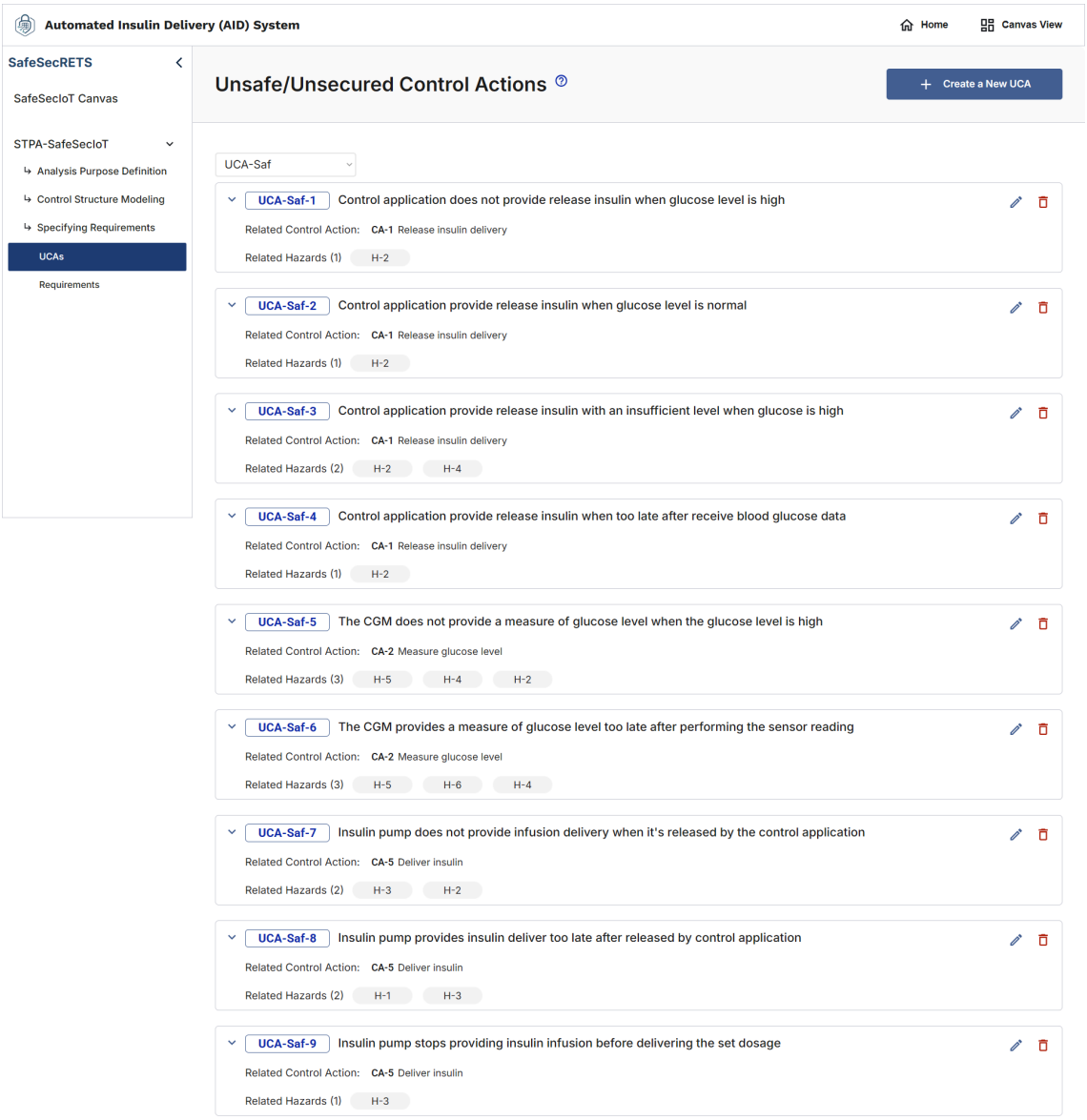


Figure 8.35: Unsafe control actions (safety).

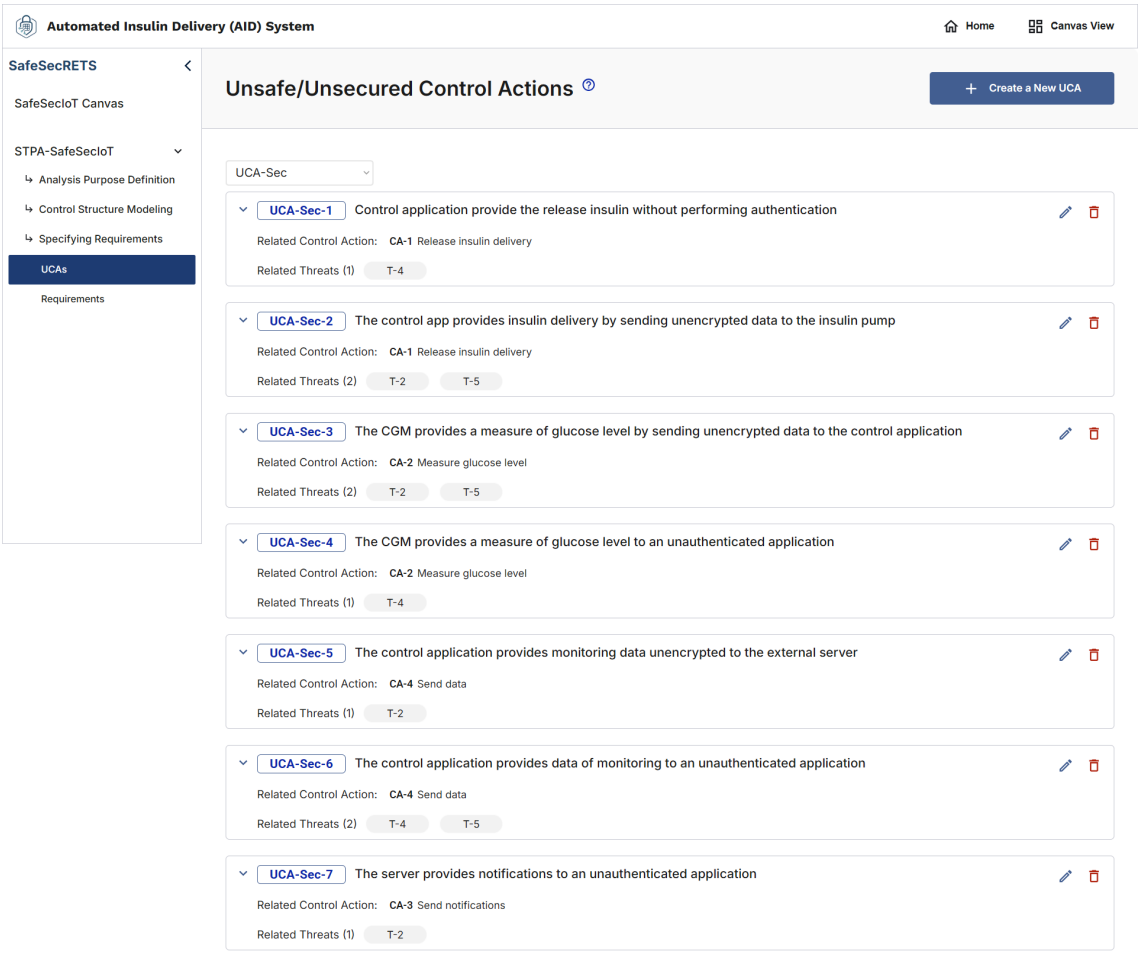


Figure 8.36: Unsecured control actions (security).

Once the system's UCAs have been defined, safety and security requirements can be specified based on them, with a view to mitigating potential vulnerabilities that could lead to hazards and threats. Each safety or security requirement includes the code and description of the related UCA and the code of the related hazards or threats, respectively. These hazards and threats must be avoided by complying with the requirements defined at the end of this analysis and specification process.

Similar to system-level safety and security constraints, safety and security requirements can also be generated with the help of the tool's AI assistant. In this case, the inputs are from the system's UCAs, which are processed by the LLM that received a prompt with specific instructions for safety and security specification based on the analysis performed, and the requirements are generated automatically from each UCA that has been defined.

Figures 8.37 and 8.38 present the lists of safety and security requirements, respectively, defined for the AID system.

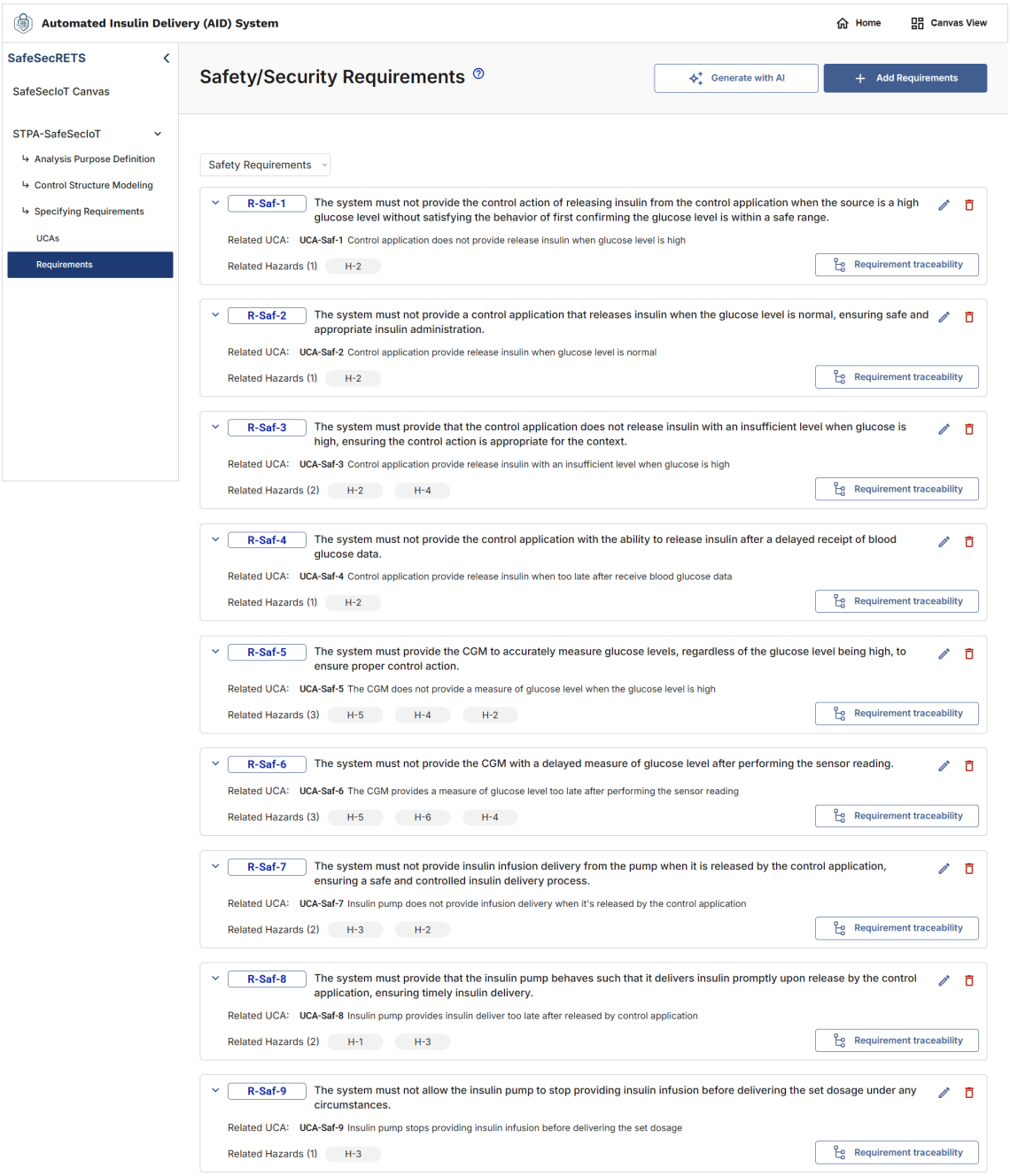


Figure 8.37: Safety requirements for the AID system.

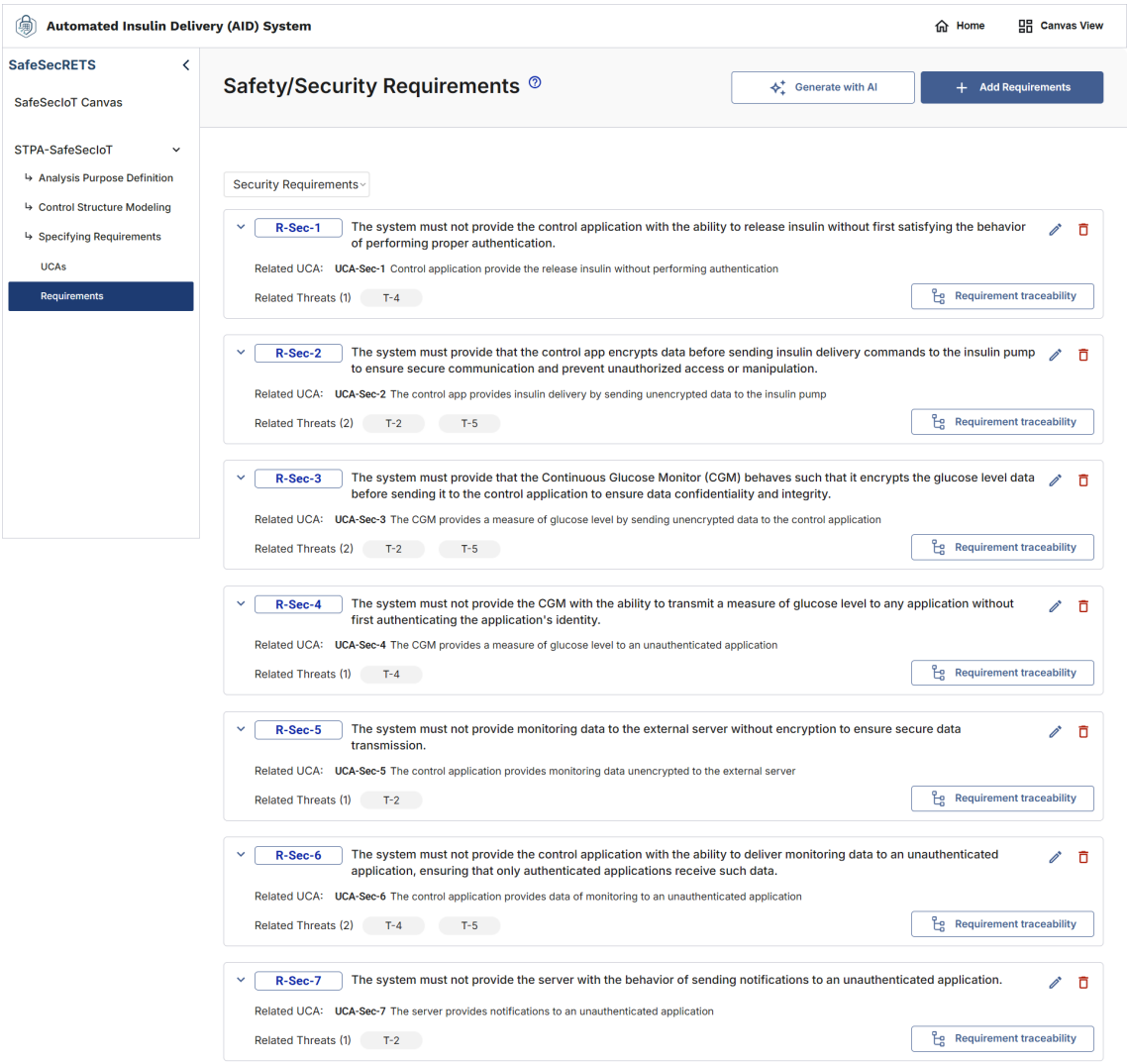


Figure 8.38: Security requirements for the AID system.

Traceability of Requirements and Analysis Information

At the end of the analysis and specification process, it is possible to generate a complete report on the traceability of the requirement, based on the STPA-based analysis process performed in the *SafeSecRETS* tool. Throughout the process, the tool establishes the link between safety or security requirements and unsafe control actions, hazards, restrictions, losses, and critical system assets. An example of a safety requirement traceability report is shown in Figure 8.39.

The analyzed requirement specifies that the AID system should not automatically release insulin in response to a high glucose level without first confirming that this value is within a safe range. As can be seen in the report, this requirement relates to UCA-Saf-1, which describes the unsafe condition of the application not delivering insulin when the glucose level is high. This condition can lead to hazard H-2, which addresses the loss of

connectivity between critical components (CGM, controller, and pump), which can lead to outdated decisions.

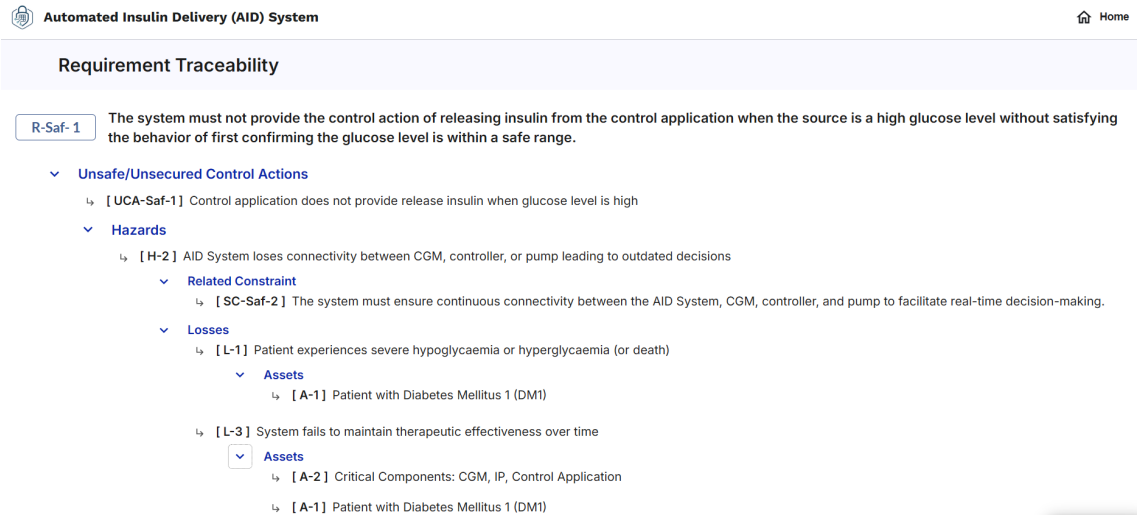


Figure 8.39: Traceability of an AID system safety requirement (part 1).

Based on hazard H-2, the SC-Saf-2 constraint is established, determining that the system must maintain continuous connectivity to ensure real-time decisions. The chain is completed by identifying losses L-1 and L-3, which represent serious consequences, such as severe hypoglycemia or hyperglycemia (or death) and the loss of therapeutic efficacy over time. These losses are directly associated with assets A-1 and A-2, which include the patient with type 1 diabetes and the critical system components. Thus, traceability ensures that each analytical element is connected, providing an integrated view that supports verification and validation of system safety.

Furthermore, the presented traceability also describes the execution flow from the control actions and the elements involved, as shown in Figures 8.40 and 8.41. Control action CA-1, responsible for triggering insulin delivery in the AID system, is initiated by Controller C-1 (control application), operating through a hardware component C-HW-3 (Android/iOS smartphone) and the software component C-SW-1 (continuous glucose monitoring application). Execution of this action involves responsibility R-7, which manages data exchange between the CGM, application, insulin pump, and cloud, and is subject to security constraint R-3, which requires user authentication before performing sensitive operations.

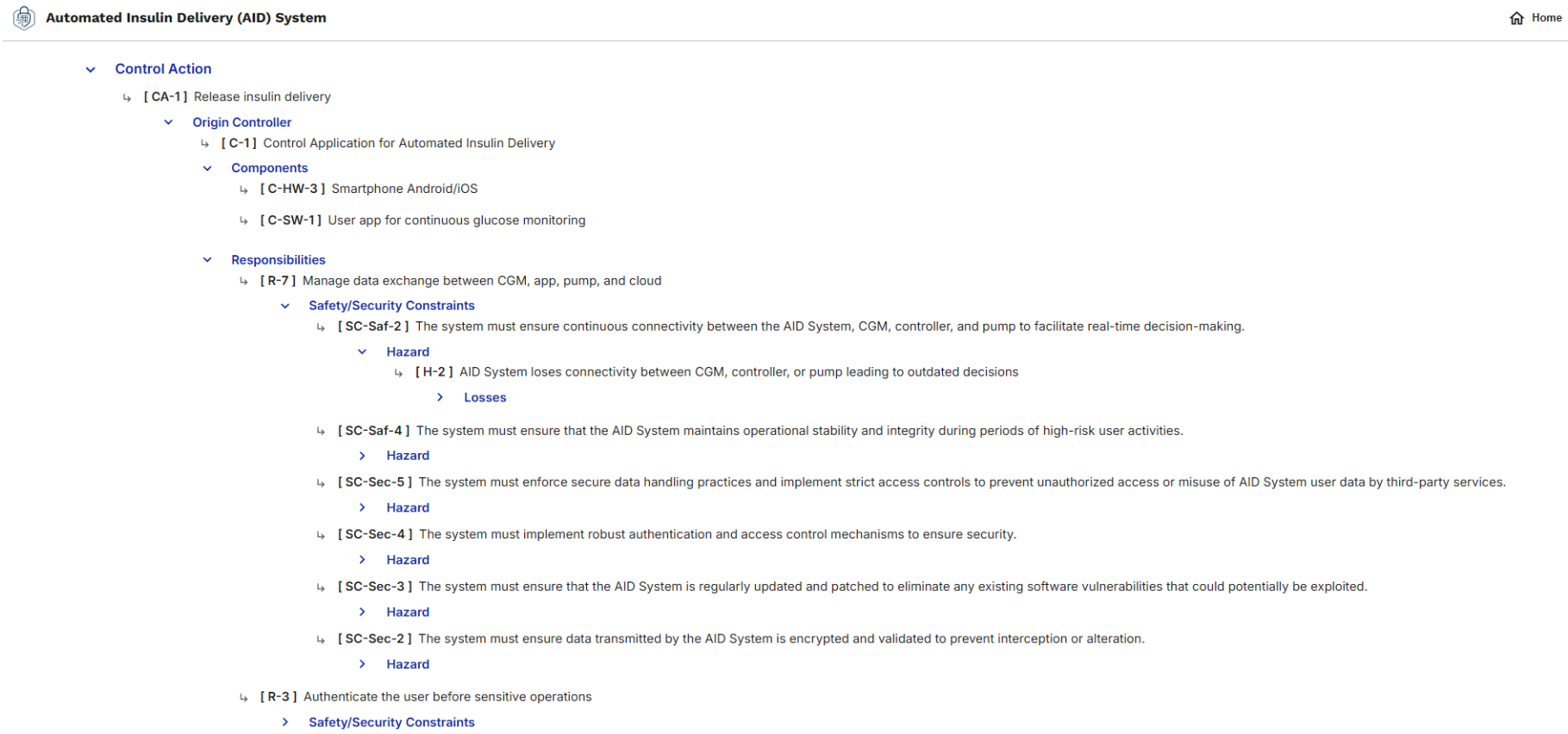


Figure 8.40: Traceability of an AID system safety requirement (part 2).

The control action has C-3 as its target controller, the insulin pump, which provides feedback F-1 to the source controller (C-1) indicating the release status. The traceability mapping in the report also identifies the sending controller (C-3) and the target controller (C-1) in the feedback flow, as well as the process-related data, which in this case refers to the administered insulin dose (bolus). This chain provides an integrated view of the control action, relating actors, responsibilities, constraints, feedback, and data, contributing to the implementation of safety and security requirements in critical automatic insulin delivery systems.

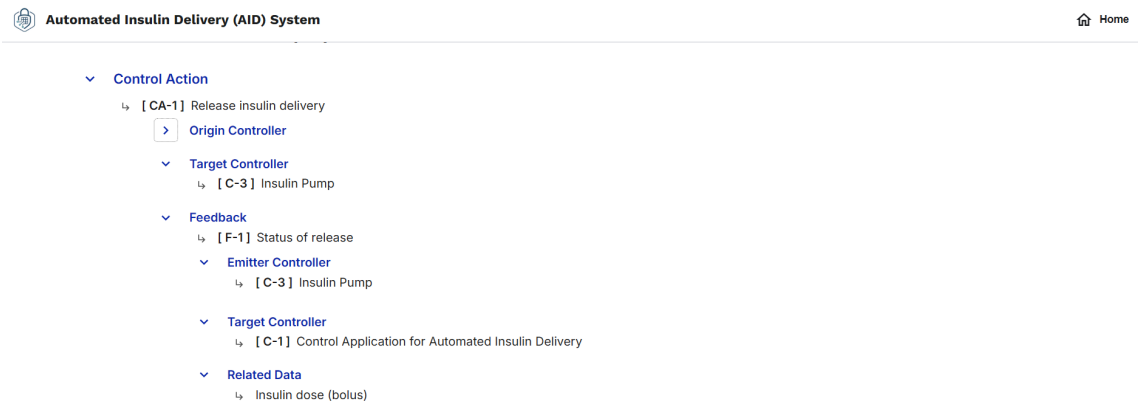


Figure 8.41: Traceability of an AID system safety requirement (part 3).

8.4 Chapter Summary

This chapter presented *SafeSecRETS*, a tool that integrates strategic planning with structured analysis of safety and security requirements for critical IoT systems. By combining the *SafeSecIoT Canvas* and *STPA-SafeSecIoT*, it supports initial scoping and requirements elicitation through a visual and systematic approach. This integration ensures traceability between planning and analysis elements, helping translate early project decisions into complete and consistent specifications.

By embedding strategic planning at the outset, *SafeSecRETS* fosters a shared understanding of project goals, context, and constraints among stakeholders. This alignment reduces ambiguities, facilitates early conflict detection between safety and security requirements, and enhances the overall quality, reliability, and informed decision-making throughout the requirements engineering process.

CHAPTER 9

Evaluation of the Safety and Security RE Process for Critical IoT Systems by Experts

In line with the empirical evaluation stage of the DSR model, this chapter presents the planning, execution, and outcomes of a study conducted with RE and safety/security experts. The purpose of the evaluation was to investigate the perceptions of professionals with practical experience in areas related to safety and security RE for critical IoT systems regarding the artifacts developed in this research: the *SafeSecIoT Canvas*, supporting project planning and requirements elicitation; the *STPA-SafeSecIoT* method, aimed at safety and security analysis and specification; and the *SafeSecRETS* tool, which integrates and operationalizes these artifacts to facilitate the RE process.

9.1 Evaluation Planning

This section describes the planning of the evaluation, which guided its conduct and subsequent analysis of the results obtained.

9.1.1 Evaluation Design

In order to evaluate the artifacts developed in the context of this research with RE experts, we adopted the survey technique [Wohlin, 2014]. A survey is a research instrument based on structured questionnaires, usually applied to a group of participants, with the aim of collecting standardized data on opinions, perceptions, behaviors, or characteristics. It serves to obtain comparable information from various respondents, allowing the identification of trends, patterns, and relationships between variables of interest in a study [Karlstrom et al., 2002].

This technique is based on selecting a sample of the population of interest and applying a questionnaire designed to obtain the information necessary for the research. The questionnaires are answered by the sample, and the information collected is then organized and can be treated quantitatively or qualitatively [Wohlin, 2014].

Considering the objects to be evaluated in this research (the artifacts developed to support the safety and security RE process of critical IoT systems), we planned a survey-based evaluation with quantitative and qualitative analyses based on the evaluation instruments and metrics defined for each instrument, as shown in Figure 9.1.

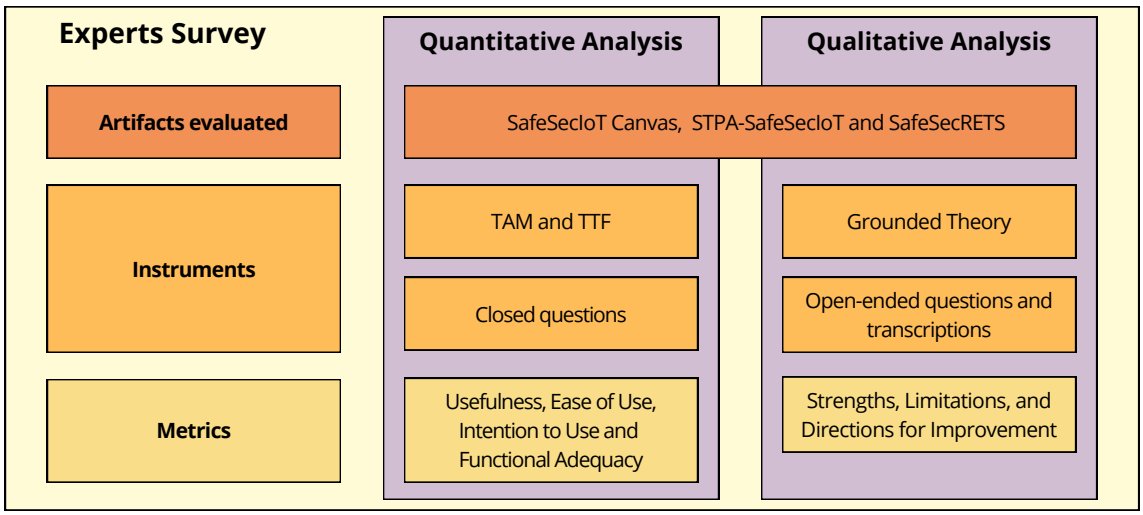


Figure 9.1: Experts evaluation design.

The evaluation was divided into two parts. The first part focused on the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts, also referred to as "conceptual artifacts" because they are respectively a canvas model and an STPA-based method. The second part aimed to evaluate the *SafeSecRETS* tool, which implements and integrates the two artifacts mentioned above in the form of software to support the safety and security RE process. Details of the planning and execution of the assessment will be presented below.

9.1.2 Evaluation Instruments

Two tools were adopted in the quantitative assessment: i) the Technology Acceptance Model (TAM) [Davis, 1989], due to its solid theoretical basis and widespread use for empirical validation in the field of technology acceptance; and ii) the Task-Technology Fit (TTF) [Goodhue and Thompson, 1995], which aims to assess the adherence of technology to user tasks/requirements. In the qualitative part, the methodological procedures of Grounded Theory (GT) [Glaser and Strauss, 1967, Corbin and Strauss, 2014] were adopted, which allows the identification of patterns, categories, and relationships from the collected data, in addition to the construction of theoretical explanations based on the data, without imposing previous models.

The quantitative analysis considered the three constructs of TAM [Davis, 1989]: i) Perceived Usefulness (PU), which measures how much the user believes that using the technology will improve their performance or bring practical benefits; ii) Perceived Ease

of Use (PEOU), which assesses how much the user believes that the technology is easy to learn and use, without excessive effort; and iii) Intention to Use (ITU), which measures the user's willingness to effectively adopt the technology in the future. These three constructs constitute a consolidated set of indicators of the acceptance of technological systems (e.g., artifacts), allowing for the evaluation of both users' perceptions and their predisposition to adopt the proposed solution [Venkatesh et al., 2003].

The qualitative analysis used the methodological procedures of GT. This methodology, proposed by [Glaser and Strauss, 1967], has several aspects that can be applied in qualitative studies. The Straussian line [Corbin and Strauss, 2014], adopted in this study, systematizes the method of data collection and analysis with well-established stages of coding, which are: open, axial, and selective.

9.1.3 Purpose of the Evaluation

The overall objective is to evaluate, with experts, the quality in use of the artifacts proposed to support the safety and security requirements engineering (RE) process in critical IoT systems. This study seeks to investigate whether, and how, the use of these artifacts contributes to the alignment between safety and security requirements, as well as to the effective conduct of the RE process. In this study, the following quality characteristics were evaluated:

- Perceived Usefulness (PU): assesses the degree to which an experienced specialist believes that the use of artifacts will increase their performance and effectiveness in safety and security RE work.
- Perceived Ease of Use (PEOU): assesses the degree of ease with which artifacts are understood and applied/used by specialists.
- Intention to Use (ITU): assesses the intention/willingness of an expert to use the artifacts in the future; verifies whether they intend to adopt or continue using a solution based on their experience, perceptions, and needs.
- Task-Technology Fit (TTF): applied only to the *SafeSecRETS* tool, assesses the degree to which a technology adequately supports the tasks that the user needs to perform; verifies whether the technology provides the necessary functionalities and capabilities for the expert to perform their activities efficiently and effectively.

9.1.4 Research Questions

To guide the empirical evaluation process based on the DSR Model, we established the following research questions:

- **RQ1 - How useful is the RE process for the safety and security of critical IoT systems, based on the proposed artifacts?** This question seeks to evaluate the Perceived Usefulness (PU) of the proposed process, whether based on the *SafeSecIoT Canvas* model in conjunction with the *STPA-SafeSecIoT* method, or supported by the *SafeSecRETS* tool, as perceived by requirements engineers who are experts in safety/security and/or IoT and/or STPA. It is addressed through the six research items for the PU construct in TAM.
- **RQ2 – How easy to use is the safety and security RE process for critical IoT systems, based on the proposed artifacts?** This question seeks to assess the Perceived Ease of Use of the proposed process, whether based on the *SafeSecIoT Canvas* model in conjunction with the *STPA-SafeSecIoT* method, or supported by the *SafeSecRETS* tool, as perceived by requirements engineers who are experts in safety/security and/or IoT and/or STPA. It is addressed through the six research items for the PEOU construct in TAM.
- **RQ3 – What is the intended use of the safety and security RE process for critical IoT systems, based on the proposed artifacts?** This question seeks to assess the Intended Use of the proposed process, whether based on the *SafeSecIoT Canvas* model in conjunction with the *STPA-SafeSecIoT* method, or supported by the *SafeSecRETS* tool, as perceived by requirements engineers who are experts in safety/security and/or IoT and/or STPA. It is addressed through the three research items for the ITU construct in TAM.
- **RQ4 – Is the proposed tool suitable for implementing the safety and security RE process for critical IoT systems, based on the proposed artifacts?** This question seeks to evaluate the Task-Technology Fit of the tool developed based on the *SafeSecIoT Canvas* model in conjunction with the *STPA-SafeSecIoT* method, as perceived by requirements engineers who are experts in safety/security and/or IoT and/or STPA. It is addressed through five TTF-based research items.
- **RQ5 – What are the strengths of the safety and security RE process for critical IoT systems based on the proposed artifacts?** This research question seeks to collect feedback on the process implemented by the artifacts from the perspective of requirements engineers, focusing on identifying the strengths and contributions of the approach. This research question is addressed through an open-ended question and transcripts of the evaluation meetings.
- **RQ6 – How could the safety and security RE process for critical IoT systems be improved?** This research question seeks to collect feedback on the process implemented by the artifacts from the perspective of requirements engineers, focusing on identifying limitations and potential improvements. This research question is addressed through an open-ended question and transcripts of the evaluation meetings.

9.1.5 Artifacts Under Evaluation and Roadmap

The evaluation of the artifacts was conducted in two sequential parts within the same meeting, following the plan outlined below:

- Part I: evaluation of the *SafeSecIoT Canvas* model for project planning and elicitation of critical IoT system requirements, and the *STPA-SafeSecIoT* method for analysis and specification of safety and security requirements. Activities:
  - Standardized video presentation of the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts and example of use;
  - Space for questions (if necessary) or comments from the expert (throughout the meeting);
  - Evaluation through closed questions based on TAM and open questions covering strengths, limitations, and suggestions for improvement (one question for each artifact and one about the integration between them).
- Part II: evaluation of the *SafeSecRETS* tool, which implements and integrates the two artifacts evaluated in part I, in the form of a software product. Activities:
  - Standardized video presentation of the *SafeSecRETS* tool and example of use;
  - Analysis of a critical IoT system project (AID system) within the tool and open interactions with it;
  - Space for questions (if necessary) or comments from the expert;
  - Evaluation through closed questions based on TAM and TTF, and open questions covering strengths, limitations, and suggestions for improvement.

Participation in the evaluation was voluntary, upon invitation to each specialist. Before the start of the evaluation sessions, the Informed Consent Form (*Termo de Consentimento Livre e Esclarecido* (TCLE)) was sent to each participating specialist via email. All participants read and signed the consent form as required by CEP/UFG, in accordance with the research project approved for this evaluation[1].

9.1.6 Selection of Context and Participants

Population

The population defined for this assessment consisted of RE experts (academics and/or professionals) with substantial experience in one or more of the following areas: safety requirements, security requirements, IoT systems, and STPA. Participation required

[1] Registered under the Certificate of Ethical Review (CAAE): 83628024.5.0000.5083

theoretical knowledge or consolidated practical experience in at least one of these areas, to ensure that respondents could provide informed and relevant evaluations.

Sample and selection of participants

To meet the defined population criteria, considering both specific areas of knowledge and their intersections, we invited researchers and professionals with recognized expertise in these domains. The study included a total of 11 experts: nine PhDs and two PhD candidates. Ten participants were professors or researchers affiliated with different universities in Brazil, with one from Portugal, and one was a professional/researcher from the Brazilian Air Force (FAB) holding a PhD in engineering and systems safety.

Figure 9.2 provides an overview of the participants' institutions, encompassing a total of ten different organizations, namely:

- Ten professors/researchers from nine universities:
  - Eight Brazilian universities: UFPE, UFC, IF Baiano, UFRJ, UNIRIO, ITA, UFJF, and Católica SC;
  - One Portuguese university: IPBeja.
- One professional/researcher from FAB

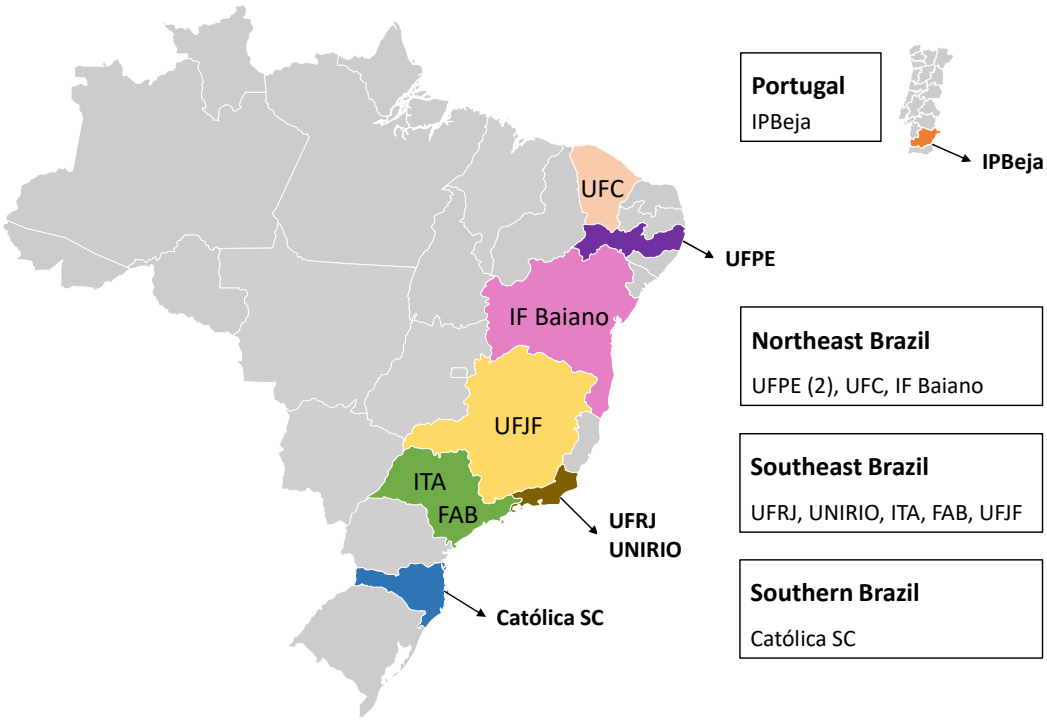


Figure 9.2: Map of institutions of the experts participating in the evaluation.

9.1.7 Profile of Experts

At the beginning of the evaluation, all experts completed a profile questionnaire designed to contextualize and qualify their responses, enabling a more accurate interpretation of the data collected on the artifacts. This questionnaire served to characterize the experts by capturing their knowledge and experience in the domains relevant to the RE activities and tasks supported by the evaluated artifacts.

Accordingly, we asked each expert to assess their level of knowledge and experience in the following areas: RE, system safety, system security, IoT, and STPA. Four levels of expertise were defined for this purpose:

1. No knowledge or experience;
2. Theoretical knowledge, without practical experience;
3. Theoretical knowledge, with some practical experience;
4. Expert, with consolidated practical experience.

Table 9.1 shows the responses of the 11 participants (self-assessment) regarding their level of knowledge/experience in each of the areas.

Table 9.1: Level of knowledge/experience in each area

Participant	ER	Safety	Security	IoT	STPA
P1	3	4	4	3	2
P2	3	3	3	2	2
P3	4	3	4	4	1
P4	4	4	2	3	4
P5	3	4	2	1	4
P6	3	2	3	3	1
P7	4	3	3	2	2
P8	4	2	3	4	1
P9	3	3	2	3	3
P10	4	4	3	3	3
P11	4	4	4	3	4

Figure 9.3 shows the radar charts (11-vertex polygon) generated from the responses presented in the previous table. A chart was generated for each area of knowledge, with one vertex representing each participating specialist. The four levels of knowledge/experience evaluated are represented as levels within the polygon of each chart.

From the responses collected, we can analyze that all participants have theoretical knowledge with practical experience in RE, with six of them being specialists with

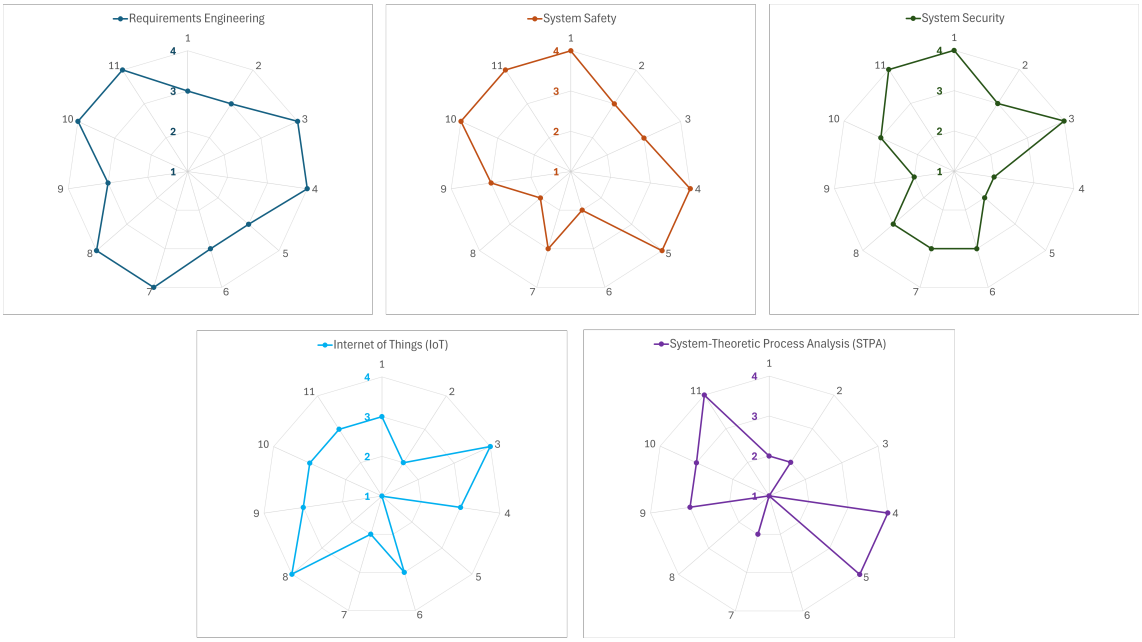


Figure 9.3: Participating experts: level of knowledge/experience declared.

consolidated practical experience. In addition, 10 of the 11 participants also have theoretical knowledge with practical experience in system safety, with five of them being specialists with consolidated practical experience.

Regarding system security, eight participants have theoretical knowledge with practical experience, three of whom are specialists with consolidated practical experience. As for IoT, eight participants also have theoretical knowledge with practical experience, two of whom are specialists with consolidated practical experience. And for STPA, five participants have theoretical knowledge with practical experience, three of whom are specialists with consolidated practical experience.

Since a participant may be an expert in more than one of the areas of knowledge presented, we conducted an analysis to understand the correlation between the areas of knowledge for the 11 participants in the assessment. The purpose of this analysis is not to generalize results for a population, but only to better understand the sample considered.

Table 9.2 shows the correlation measures between the areas of knowledge for the sample considered. Based on the results, it can be stated that, for the 11 specialists, there is a strong correlation between the areas of STPA and system safety (positive correlation of 0.8). There is also some correlation, albeit weaker, between the areas of RE and IoT knowledge and also between IoT and system security (positive correlations of 0.45 and 0.44, respectively). In addition, there is some correlation between the areas of IoT and system safety, as well as between the areas of STPA and system security (negative correlations of 0.35 and 0.42, respectively).

Table 9.2: Participating experts: correlation among areas of expertise.

	Requirements Engineering	**System Safety**	**System Security**	**Internet of Things**	**STPA**
RE	—	—	—	—	—
System Safety	0.089	—	—	—	—
System Security	0.247	0.000	—	—	—
Internet of Things	0.458	-0.357	0.443	—	—
STPA	0.043	0.800**	-0.426	-0.480	—

Note. * $p < .05$, * $p < .01$, ** $p < .001$.

9.2 Quantitative Analysis: Evaluation Questionnaires

Below we present the quantitative analysis process performed based on the application of TAM and TTF to the respective artifacts evaluated. The analysis reported aims to answer research questions **RQ1 to RQ4** presented in subsection 9.1.4.

9.2.1 Reliability of questionnaires

In order to ensure the internal consistency of the instruments used in the evaluation of the artifacts and the tool developed, Cronbach's alpha coefficient was applied to the questionnaires based on the TAM and TTF models. Considering that the items in these questionnaires were organized into theoretical constructs such as PU, PEOU, and ITU (in the case of TAM, for example), it is important to verify whether the items in each construct presented homogeneity in the responses provided by the experts.

The application of Cronbach's alpha allows us to assess the degree of correlation between the items that make up each dimension of the model, giving greater statistical robustness to the analysis. Thus, the use of this coefficient is justified by the need to ensure that the questionnaires used are reliable and adequate to capture the experts' perceptions regarding the acceptance and adequacy of the proposed solutions.

SafeSecIoT Canvas model and *STPA-SafeSecIoT* method

Table 9.3 presents data on the internal consistency of the instruments used in the evaluation of the *SafeSecIoT Canvas* model and the *STPA-SafeSecIoT* method. Considering the Cronbach's alpha values found (between 0.728 and 0.890), it is possible to state that all questionnaires have good internal consistency, which indicates that the items in each construct are consistent with each other and measure the same concept in a cohesive manner.

The *SafeSecIoT Canvas* model showed high average values in all dimensions (PU = 6.20; PEOU = 6.24; ITU = 6.37), on a scale of 1 to 7, with good internal consistency

Table 9.3: Summary metrics by artifact and construct (mean, SD, and Cronbach's α).

Artifact	Construct	Mean	Standard deviation	Cronbach's α
SafeSecIoT Canvas	Perceived Usefulness (PU)	6.20	0.628	0.827
	Perceived Ease of Use (PEOU)	6.24	0.698	0.844
	Intention to Use (ITU)	6.37	0.564	0.728
STPA–SafeSecIoT	Perceived Usefulness (PU)	5.98	0.489	0.751
	Perceived Ease of Use (PEOU)	5.89	0.589	0.869
	Intention to Use (ITU)	6.48	0.603	0.890

(alpha > 0.7 in all cases). This indicates that experts perceived the artifact as useful, easy to use, and with a high intention to adopt, although the reliability of the intention to use scale (alpha = 0.728) is slightly lower than the others.

The *STPA-SafeSecIoT* method obtained slightly lower averages in PU (5.98) and PEOU (5.89) compared to the *SafeSecIoT Canvas*, but an even higher ITU (6.48), with excellent internal consistency ($\alpha = 0.890$). These results suggest that, although the method may require greater cognitive effort or be perceived as more complex, participants highly value it for its practical applicability, which is reflected in a strong intention to use it.

SafeSecRETS tool

Table 9.4 shows the data relating to the internal consistency of the instruments used in the evaluation of the *SafeSecRETS* tool. Based on the Cronbach's α values presented (above 0.90), it can be said that all questionnaires have high internal consistency.

Table 9.4: Summary metrics by construct (mean, SD, and Cronbach's α)

Construct	Mean	Standard deviation	Cronbach's α
Perceived Usefulness (PU)	6.55	0.606	0.905
Perceived Ease of Use (PEOU)	6.59	0.643	0.930
Intention to Use (ITU)	6.70	0.690	0.915
Task–Technology Fit (TTF)	6.18	1.20	0.953

The results confirm that the instruments used were statistically reliable, legitimizing the analyses performed based on the experts' perceptions of the evaluated tool. In addition, the high averages (all above 6.0 on a scale of 1 to 7) indicate a consistent positive perception of the evaluated tool by the experts. More details of this analysis will be presented in the next section, which details the descriptive statistics.

9.2.2 Descriptive statistics and analysis based on the TAM model

Part 1. *SafeSecIoT Canvas* model and *STPA-SafeSecIoT* method

For this stage of the analysis, we used the general statistics (mean, standard deviation, and Cronbach's alpha) presented in Table 9.3 (i) and the detailed descriptive statistics for each questionnaire, which will be presented throughout the analysis of each construct.

Perceived Usefulness (PU-TAM) of the *SafeSecIoT Canvas* model. The high mean (6.20 on a scale of 1 to 7) indicates that experts agree that the *SafeSecIoT Canvas* model is useful for planning critical IoT system projects. The standard deviation is low (0.628), demonstrating a high degree of agreement among the evaluators and suggesting a uniform perception. Cronbach's alpha of 0.827 demonstrates excellent internal consistency, which validates that the questionnaire items used to measure PU are well aligned and cohesive in capturing the construct. Table 9.5 presents the detailed descriptive statistics for PU-TAM of the *SafeSecIoT Canvas* model.

Table 9.5: Descriptive statistics for PU of the *SafeSecIoT Canvas* model.

	PU-TAM-1	**PU-TAM-2**	**PU-TAM-3**	**PU-TAM-4**	**PU-TAM-5**	**PU-TAM-6**
N	9	9	9	9	9	9
Missing	0	0	0	0	0	0
Mean	6.44	6.22	6.33	5.67	6.33	6.22
Median	6	6	6	6	7	6
Mode	6.00	6.00	6.00	7.00	7.00	7.00
Standard deviation	0.527	0.972	0.500	1.22	0.866	0.833
Minimum	6	4	6	4	5	5
Maximum	7	7	7	7	7	7

Ease of Use (PEOU-TAM) of the *SafeSecIoT Canvas* model. With an average slightly higher than that of PU-TAM (6.24), participating experts rate the model as easy to use. The standard deviation is slightly higher (0.698), but pointing to a consensus among the responses. Cronbach's alpha (0.844) corroborates the reliability of the questionnaire, showing that the PEOU items consistently measure the same concept. Figure 9.6 presents the detailed descriptive statistics for PEOU-TAM of the *SafeSecIoT Canvas* model.

The median and mode were between 6 and 7 for all items (6 questions) of the PU and PEOU constructs (considering the Likert scale from 1 to 7), showing a high concentration of agreement ratings on PU and PEOU for most experts, as shown in Appendix B, Section B.2. In addition, regarding PU, most of the experts' responses were in the positive range (ratings between 5 and 7). Only three responses were in the neutral

Table 9.6: Descriptive statistics for PEOU of the *SafeSecIoT Canvas* model.

	PEOU-TAM-1	**PEOU-TAM-2**	**PEOU-TAM-3**	**PEOU-TAM-4**	**PEOU-TAM-5**	**PEOU-TAM-6**
N	9	9	9	9	9	9
Missing	0	0	0	0	0	0
Mean	6.11	6.22	6.33	6.33	6.22	6.22
Median	6	6	7	6	7	6
Mode	6.00	6.00	7.00	6.00	7.00	7.00
Standard deviation	1.27	0.667	1.00	0.707	0.972	0.833
Minimum	3	5	4	5	5	5
Maximum	7	7	7	7	7	7

range, one regarding performance improvement and two regarding work effectiveness improvement. No responses were in the negative range (ratings between 1 and 3) for PU.

Regarding PEOU, most responses were also positive, with only one response in the neutral range and one in the negative range, as shown in Figure B.1 (Appendix B). Thus, according to the quantitative data, we conclude that the *SafeSecIoT Canvas* model is perceived as highly useful and is considered easy to use, consistently among experts.

Intention to Use (ITU-TAM) of the *SafeSecIoT Canvas* model. This construct obtained the highest average among the constructs for the proposed canvas model (6.37), indicating a high intention to use the artifact. According to the statistics presented in the Table 9.7, the median and mode for all items in the construct were high (6 and 7), showing a high level of agreement among experts on the intent to use the artifact.

Table 9.7: Descriptive statistics for ITU of the *SafeSecIoT Canvas* model.

	ITU-TAM-1	**ITU-TAM-2**	**ITU-TAM-3**
N	9	9	9
Missing	0	0	0
Mean	6.33	6.56	6.22
Median	6	7	6
Mode	6.00	7.00	7.00
Standard deviation	0.500	0.726	0.833
Minimum	6	5	5
Maximum	7	7	7

Perceived Usefulness (PU-TAM) of the *STPA-SafeSecIoT* method. The average (5.98 on a scale of 1 to 7) indicates that experts consider the *STPA-SafeSecIoT* method useful for the analysis and specification of safety and security in critical IoT systems. The standard deviation is very low (0.489), demonstrating a high degree of agreement among evaluators and a uniform perception of PU. Cronbach's alpha of 0.751 demonstrates good internal consistency, which validates that the questionnaire items used to measure PU are

aligned and cohesive in capturing the construct. Table 9.8 presents the detailed descriptive statistics for PU-TAM of the *STPA-SafeSecIoT* method.

Table 9.8: Descriptive statistics for PU of the *STPA-SafeSecIoT* method.

	PU-TAM-1	**PU-TAM-2**	**PU-TAM-3**	**PU-TAM-4**	**PU-TAM-5**	**PU-TAM-6**
N	9	9	9	9	9	9
Missing	0	0	0	0	0	0
Mean	6.11	6.00	6.22	5.56	5.78	6.22
Median	6	6	6	6	6	6
Mode	6.00	6.00	6.00	6.00	6.00	6.00
Standard deviation	0.333	0.866	0.667	1.01	0.667	0.667
Minimum	6	4	5	4	5	5
Maximum	7	7	7	7	7	7

Ease of Use (PEOU-TAM) of the *STPA-SafeSecIoT* method. With an average score practically equal to that of PU-TAM (5.89), experts have a consistent perception that the method is easy to use. The standard deviation is very low (0.589), corroborating the regularity among the responses. Cronbach's alpha (0.869) indicates high reliability of the questionnaire, showing that the PEOU items consistently measure this construct. Table 9.9 presents detailed descriptive statistics for PEOU-TAM of the *STPA-SafeSecIoT* method.

Table 9.9: Descriptive statistics for PEOU of the *STPA-SafeSecIoT* method.

	PEOU-TAM-1	**PEOU-TAM-2**	**PEOU-TAM-3**	**PEOU-TAM-4**	**PEOU-TAM-5**	**PEOU-TAM-6**
N	9	9	9	9	9	9
Missing	0	0	0	0	0	0
Mean	5.78	6.11	6.00	5.56	5.89	6.00
Median	6	6	6	6	6	6
Mode	6.00	6.00	6.00	6.00	6.00	6.00
Standard deviation	1.20	0.601	0.500	0.882	0.601	0.500
Minimum	3	5	5	4	5	5
Maximum	7	7	7	7	7	7

The median and mode were 6 for all items (6 questions) of the PU-TAM and PEOU-TAM constructs (considering the Likert scale from 1 to 7), showing high agreement on PU and PEOU for most experts, as shown in Figure B.3 (Appendix B). Regarding PU, most of the experts' responses were in the positive range (ratings between 5 and 7). Only three responses were in the neutral range, one regarding performance improvement and two regarding work effectiveness improvement. No responses were in the negative range (ratings between 1 and 3) for PU.

Regarding PEOU, most responses were also positive, with only one response in the neutral range and one in the negative range, as shown in Figure B.3 (Appendix B). Thus, according to the quantitative data, we conclude that the *STPA-SafeSecIoT* method is considered useful and easy to use, consistently among experts.

Intention to Use (ITU-TAM) of the *STPA-SafeSecIoT* method. This construct obtained a considerably higher average among the constructs that evaluated the extent of the proposed STPA method (6.48), indicating a high intention to use the artifact even though it is more complex, considering the PEOU. The ITU perceived by experts for the *STPA-SafeSecIoT* method was even higher than that observed for the canvas model, indicating that the artifact has an important relevance for the RE process of critical IoT systems, which is justified by the analysis and specification steps it supports.

According to the statistics presented in the Table 9.10, the median and mode for all items in the construct were high (6 and 7), showing a high level of agreement among experts on their intention to use the artifact.

Table 9.10: Descriptive statistics for ITU of the *STPA-SafeSecIoT* method.

	ITU-TAM-1	**ITU-TAM-2**	**ITU-TAM-3**
N	9	9	9
Missing	0	0	0
Mean	6.44	6.44	6.56
Median	7	6	7
Mode	7.00	6.00	7.00
Standard deviation	0.726	0.527	0.726
Minimum	5	6	5
Maximum	7	7	7

Part 2. *SafeSecRETS* tool

For this analysis, we used the general statistics (mean, standard deviation, and Cronbach's alpha) presented in Figure 9.4 (i) and the detailed descriptive statistics for each questionnaire, which will be presented throughout the analysis of each construct.

Perceived Usefulness (PU-TAM) of the *SafeSecRETS* tool. The high mean (6.55) indicates that the experts strongly agreed that the tool is useful for performing their tasks. The low standard deviation (0.606) demonstrates a high degree of agreement among the evaluators, suggesting a fairly uniform perception. Cronbach's alpha of 0.905 demonstrates excellent internal consistency, which validates that the questionnaire items used to measure PU are well aligned and cohesive in capturing the construct.

Table 9.11 presents detailed descriptive statistics for PU-TAM of the *SafeSecRETS* tool.

Table 9.11: Descriptive statistics for PU of the *SafeSecRETS* tool.

	PU-TAM-1	**PU-TAM-2**	**PU-TAM-3**	**PU-TAM-4**	**PU-TAM-5**	**PU-TAM-6**
N	11	11	11	11	11	11
Missing	0	0	0	0	0	0
Mean	6.55	6.64	6.45	6.27	6.55	6.82
Median	7	7	7	7	7	7
Mode	7.00	7.00	7.00	7.00	7.00	7.00
Standard deviation	0.688	0.505	0.820	1.10	0.688	0.405
Minimum	5	6	5	4	5	6
Maximum	7	7	7	7	7	7

The median and mode were 7 for all items (6 questions) of the PU-TAM construct (considering the Likert scale from 1 to 7), showing a high concentration of "strongly agree" ratings for most experts, as shown in Figure B.5 (Appendix B). In addition, all expert responses, except for one, were in the positive range (ratings between 5 and 7). Only one expert neither agreed nor disagreed that the tool would improve their work efficiency, responding in the neutral range (rating 4). No responses were in the negative range (ratings between 1 and 3). Thus, based on the quantitative data, we conclude that the tool is consistently perceived as highly useful among experts.

Ease of Use (PEOU-TAM) of the *SafeSecRETS* tool. With an average slightly higher than that of PU-TAM (6.59), participating experts strongly agree that the tool is easy to use. The low standard deviation (0.643) again points to a high consensus among responses. Cronbach's alpha (0.930) confirms the strong reliability of the questionnaire, showing that the PEOU items consistently measure the same concept.

Table 9.12 presents the detailed descriptive statistics for PEOU-TAM of the *SafeSecRETS* tool.

Table 9.12: Descriptive statistics for PEOU of the *SafeSecRETS* tool.

	PEOU-TAM-1	**PEOU-TAM-2**	**PEOU-TAM-3**	**PEOU-TAM-4**	**PEOU-TAM-5**	**PEOU-TAM-6**
N	11	11	11	11	11	11
Missing	0	0	0	0	0	0
Mean	6.45	6.55	6.55	6.64	6.64	6.73
Median	7	7	7	7	7	7
Mode	7.00	7.00	7.00	7.00	7.00	7.00
Standard deviation	1.21	0.688	0.688	0.674	0.505	0.467
Minimum	3	5	5	5	6	6
Maximum	7	7	7	7	7	7

The median and mode were 7 for all questions, indicating that most responses

were positive and that there was a high concentration of "strongly agree" ratings, as shown in Figure B.5 (Appendix B). Only one response was negative (3), regarding the ease of learning to use the tool. All other responses were in the positive range (5 to 7). Thus, based on the data evaluated, we conclude that the tool was perceived as very easy to use and that this perception is consistent among the participating experts.

Intention to Use (ITU-TAM) of the *SafeSecRETS* tool. This construct had the highest mean (6.70) among all those evaluated for the *SafeSecRETS* tool, suggesting that experts have a strong intention to adopt or recommend the use of the tool. The standard deviation was also low (0.690), indicating good agreement among participants. The alpha value (0.915) indicates that there is excellent internal consistency in the questionnaire items, reinforcing the reliability of the measurement of intention to use for experts.

Table 9.13 presents the detailed descriptive statistics for ITU-TAM of the *SafeSecRETS* tool. The median and mode were 7 for all items in the construct, indicating that most responses were "strongly agree." Only one response was in the neutral range (4), with all others in the positive range (5 to 7) and none in the negative range. Thus, based on the quantitative data, we conclude that the intention to use the tool is very high and uniform, supporting the acceptance of the proposed solution.

Table 9.13: Descriptive statistics for ITU of the *SafeSecRETS* tool.

	ITU-TAM-1	**ITU-TAM-2**	**ITU-TAM-3**
N	11	11	11
Missing	0	0	0
Mean	6.82	6.64	6.64
Median	7	7	7
Mode	7.00	7.00	7.00
Standard deviation	0.405	0.924	0.809
Minimum	6	4	5
Maximum	7	7	7

Task-Technology Fit (TTF) of the *SafeSecRETS* tool. Despite having the lowest average among the constructs evaluated (6.18), the value is still quite high (considering that the maximum value would be 7). This indicates that experts perceive that there is a good alignment between the tool's functionalities and the needs of the tasks performed. The standard deviation of 1.20 is the highest among the constructs that make up the tool evaluation, suggesting greater variation in individual perceptions in this construct. Still, Cronbach's alpha of 0.953 points to excellent consistency among the items, which validates the measured construct. Table 9.14 presents the statistics obtained in the TTF evaluation:

Even with a higher standard deviation than the other constructs, the median and mode were also 7, demonstrating strong agreement among most of the experts who

Table 9.14: Descriptive statistics for the TTF construct of the *SafeSecRETS* tool.

	TTF-1	**TTF-2**	**TTF-3**	**TTF-4**	**TTF-5**
N	11	11	11	11	11
Missing	0	0	0	0	0
Mean	6.27	6.18	6.00	6.09	6.36
Median	7	7	7	7	7
Mode	7.00	7.00	7.00	7.00	7.00
Standard deviation	1.01	1.40	1.34	1.64	1.03
Minimum	4	3	3	2	4
Maximum	7	7	7	7	7

evaluated the tool and that most responses were concentrated in the positive range (5 to 7). As shown in Figure B.6 (Appendix B), there were only three responses in the negative range, referring to: i) the tool's assistance in performing tasks with greater precision and consistency; ii) easy access to the information and elements necessary for the task; and iii) the correct implementation of the functionalities necessary to perform the tasks. Thus, we conclude that the TTF is perceived as positive, even with more pronounced variations. Thus, we conclude that the TTF is perceived as positive, even with more pronounced variations, due to differences in the profile, context, or experience of the experts, as evidenced by the profile questionnaire.

9.2.3 Comparative Analysis Between the Artifacts and the Tool

This section presents a comparative analysis between the evaluated artifacts, considering: i) the *SafeSecIoT Canvas* model together with the *STPA-SafeSecIoT* method; and ii) the *SafeSecRETS* tool. This analysis was conducted based on the averages of the responses to the TAM-based questionnaire, considering PU, PEOU, and ITU. The objective is to understand whether the tool contributes to the effectiveness and intention to use the proposed approach, in comparison to the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts.

Figure 9.4 presents a comparison between each of the questions related to PU for the artifacts (we will use the term artifacts to refer to *SafeSecIoT Canvas* and *STPA-SafeSecIoT* in this section) and the *SafeSecRETS* tool. The average difference calculated between the artifacts and the tool was 7.55%. It can be seen in the figure that the tool obtained a higher score for all questions, in relation to the average score of the artifacts, which indicates a perception of greater usefulness of the tool in relation to the artifacts, for the participating experts.

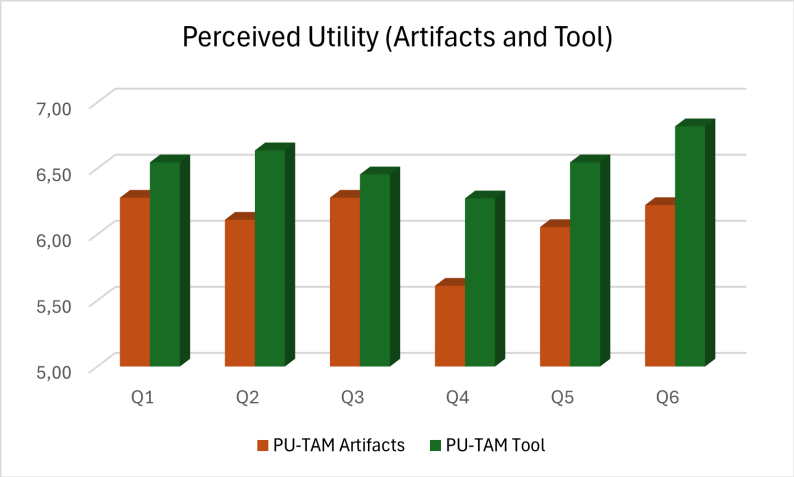


Figure 9.4: Comparison between the PU of the artifacts and the tool.

Figure 9.5 presents a similar comparison for questions related to PEOU considering the artifacts and the tool. The average difference calculated between the artifacts and the tool was 8.75%. It can be observed that the tool obtained a higher score for all questions compared to the artifacts, which indicates a perception of greater ease of use of the tool compared to the artifacts for the participating experts.

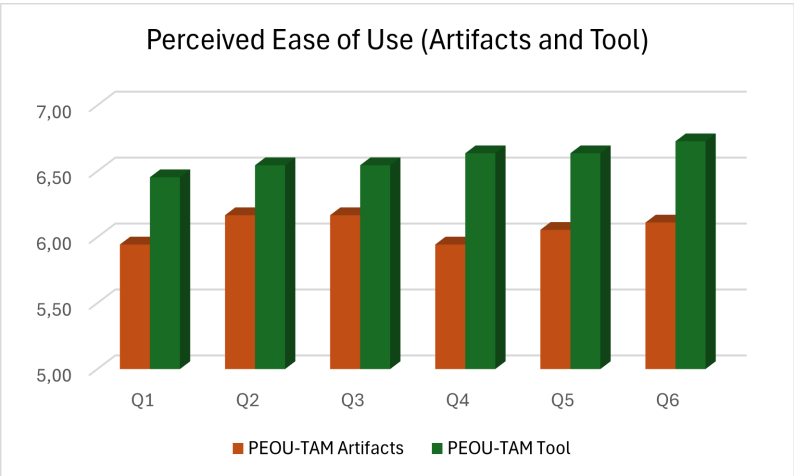


Figure 9.5: Comparison between the PEOU of the artifacts and the tool.

Figure 9.6 shows the comparison between questions regarding the intention to use the artifacts and the tool. The average difference calculated between the artifacts and the tool was 4.20%. It can be observed that, in this case as well, the tool obtained a higher score for all questions compared to the artifacts. Thus, the evaluation indicates a perception of greater intention to use the tool compared to the artifacts among the participating experts.

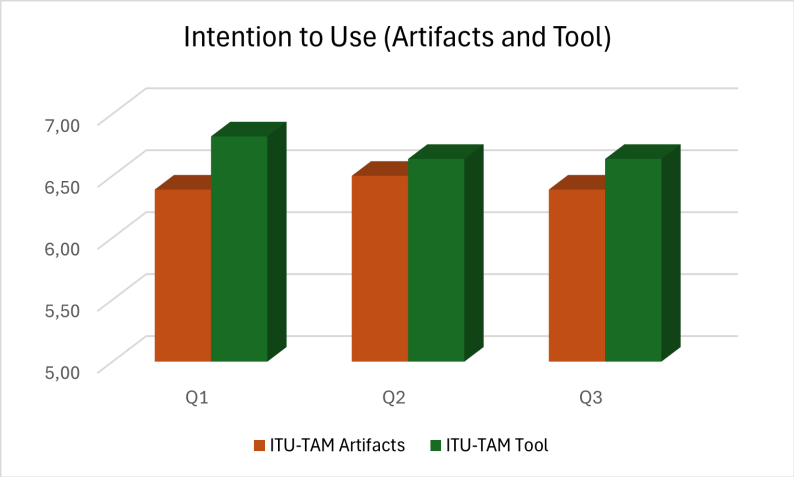


Figure 9.6: Comparison between the ITU of the artifacts and the tool.

Finally, Figure 9.7 shows the overall averages obtained for the TAM constructs, both for the artifacts and for the tool. The averages obtained indicate high levels of acceptance for both the artifacts and the tool. However, it can be observed that the tool presented higher averages in all dimensions, especially in Intention to Use, where it obtained the highest average among the constructs (6.70).

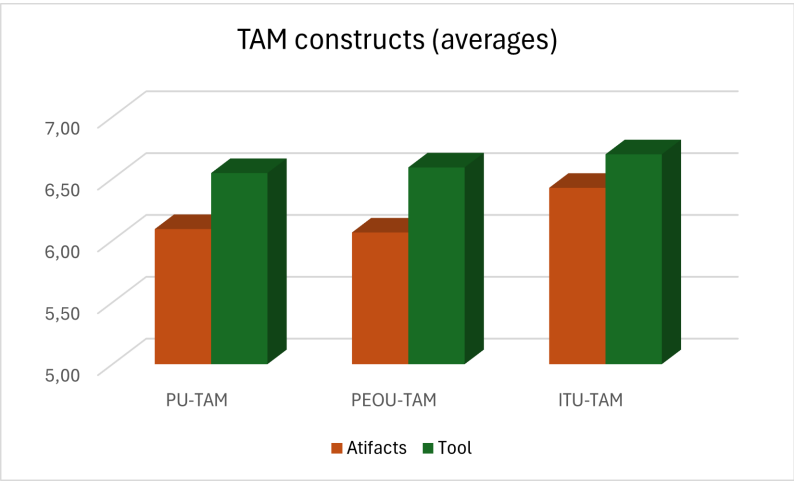


Figure 9.7: Comparison between TAM constructs for artifacts and the tool.

The differences found between the evaluations of the artifacts and the tool suggest that integrating the artifacts into a tool (software product) contributes to a more positive experience, making the approach more accessible and increasing the likelihood that experts will use it in their practices.

Considering the Likert scale from 1 to 7 used in the evaluation, both the artifacts and the tool presented high averages in all dimensions of the TAM model, indicating an overall positive perception by the experts. Even so, we can say that the tool stood

out consistently, with higher averages in all dimensions compared to the artifacts: PU (+7.55%), PEOU (+8.75%), and ITU (+4.20%).

The dispersion of the experts' evaluations was slightly higher for artifacts compared to the tool, especially in the PU and PEOU dimensions. This result may indicate that, while the tool presents a more homogeneous and accessible user experience, the artifacts require greater interpretation and prior mastery of the concepts in order for their capabilities to be fully exploited.

The results obtained indicate that:

- Both the artifacts and the tool received good evaluations from the experts in terms of usefulness, ease of use, and intention to adopt.
- PU: The tool contributes to making the process easier and more useful, in the experts' view, which can increase the efficiency of the approach.
- PEOU: Experts perceived the tool as more accessible and easier to use than the original artifacts.
- ITU: The tool is more likely to be adopted in a real-world scenario compared to the artifacts.

Thus, we can affirm that, even if there were no tool, the results of the artifact evaluation indicate that they would be useful and easy to use by RE specialists, with or without experience in STPA. However, the tool enhances the usefulness and ease of use of the proposed approach, consequently increasing its intended use (for the group of participating specialists).

9.2.4 Influence of Expert Profile

To examine whether participants' profiles influenced their perception of artifact quality, we conducted a complementary analysis based on the profile questionnaire, in which each expert self-assessed their knowledge in the areas of safety, security, IoT, and STPA. Knowledge/experience was rated on a four-level scale (Table 9.1), with 1 indicating no knowledge or experience and 4 indicating practical expertise.

For the analysis in this section, participants were divided into two groups according to their responses: Group 1, experts who reported a knowledge level of 1 or 2 in any of the areas; and Group 2, experts who reported a level of 3 or 4. Since all participants indicated a level of 3 or 4 in RE (and also in safety and/or security), requirements engineering was taken as a shared baseline, allowing us to state that all participants are experts in RE for critical systems.

Table 9.15 shows the averages for each construct, considering the knowledge and experience declared by each participating expert in the areas of knowledge. Below we present an analysis of the results obtained.

Table 9.15: Table of averages for each construct and by group, considering declared knowledge and experience.

Artifact	**Construct**	**Overall Average**	**Safety**		**Security**		**STPA**		**IoT**	
			Group 1	**Group 2**	**Group 1**	**Group 2**	**Group 1**	**Group 2**	**Group 1**	**Group 2**
	PU	6,20	6,44	6,08	6,28	6,17	6,50	5,97	6,44	6,08
SafeSecIoT	PEOU	6,24	6,50	6,11	6,33	6,19	6,63	5,93	6,72	6,00
Canvas	ITU	6,37	6,78	6,17	6,11	6,50	6,83	6,00	6,67	6,22
	TAM	**6,27**	**6,57**	**6,12**	**6,24**	**6,29**	**6,65**	**5,97**	**6,61**	**6,10**
	PU	5,98	6,22	5,86	6,17	5,89	6,13	5,87	6,17	5,89
STPA–	PEOU	5,89	5,83	5,92	6,22	5,72	5,88	5,90	6,39	5,81
SafeSecIoT	ITU	6,48	6,67	6,39	6,44	6,50	6,75	6,27	6,67	6,22
	TAM	**6,12**	**6,24**	**6,06**	**6,28**	**6,04**	**6,25**	**6,01**	**6,41**	**5,97**
	PU	6,55	7,00	6,44	6,94	6,40	6,42	6,70	6,11	6,71
	PEOU	6,59	6,92	6,52	6,83	6,50	6,67	6,50	6,50	6,63
SafeSecRETS	ITU	6,70	7,00	6,63	7,00	6,58	6,67	6,73	7,00	6,58
	TAM	**6,61**	**6,97**	**6,53**	**6,93**	**6,49**	**6,58**	**6,64**	**6,54**	**6,64**
	TTF	**6,18**	**6,90**	**6,02**	**5,95**	**6,80**	**6,17**	**6,20**	**6,18**	**6,20**

Part 1. *SafeSecIoT Canvas* Model and *STPA-SafeSecIoT* Method

The analysis of the TAM construct averages, considering experts' self-declared knowledge in Safety, Security, IoT, and STPA, revealed a consistent pattern in the evaluation of the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts. Overall, mean scores were high across all constructs, both for the full sample and for the groups defined by knowledge area. The lowest mean was 5.72 (on a 1–7 scale), observed for PEOU by Group 2 in the evaluation of *STPA-SafeSecIoT*, while most means ranged between 6 and 7. Furthermore, in almost all cases, Group 1 participants assigned slightly higher scores than those in Group 2, with only a few exceptions to this trend.

Analyzing the *SafeSecIoT Canvas* assessments, Group 1 generally reported higher average scores across almost all constructs compared to Group 2. For instance, in PU, scores were 6.44 vs. 6.08 in Safety, 6.28 vs. 6.17 in Security, 6.50 vs. 5.97 in STPA, and 6.44 vs. 6.08 in IoT. This suggests that experts with less familiarity perceive greater utility in the artifacts. A similar pattern appears in PEOU, with Group 1 consistently achieving higher averages, reaching 6.72 in IoT compared to 6.00 for Group 2. In ITU, Group 1 again stood out (6.78 in Safety, 6.72 in IoT), indicating a strong intention to use the artifacts.

For *STPA-SafeSecIoT*, the pattern was very similar, with Group 1 generally reporting higher scores across PU, PEOU, and ITU. In PU, scores ranged from 6.22 to 6.50 for Group 1 and 5.80 to 5.89 for Group 2; in PEOU, from 6.33 to 6.72 versus 5.72 to 6.00; and in ITU, from 6.44 to 6.75 versus 5.50 to 6.27, respectively. These results indicate that, while both groups recognize the artifact's value, participants with less experience perceive a greater practical benefit.

This evaluation pattern suggests that Group 2 participants (those with more experience in the respective knowledge areas) tend to be more critical, particularly regarding *STPA-SafeSecIoT*, likely because they are already familiar with the concepts and practices that the method seeks to structure. In contrast, Group 1 participants, with less prior experience, seem to benefit more from the structured guidance provided by both artifacts, which offer a detailed framework to support safety and security RE activities in critical IoT systems.

Overall, the results indicate that the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts have strong adoption potential. They are valuable for multidisciplinary teams or analysts with limited experience, while still being recognized as relevant by experts, who, although more critical in their assessments of PU and PEOU, acknowledge their practical utility in guiding the RE process for critical IoT projects.

Part 2. *SafeSecRETS* Tool

In the case of the *SafeSecRETS* tool, the differences between groups also favored Group 1, though with some nuances across constructs. In PU, for instance, Group 1 achieved averages of 7.00 (Safety) and 6.94 (Security), compared to 6.44 and 6.40 for Group 2. A similar pattern appears in PEOU, with Group 1 scoring 6.92 (Safety) and 6.33 (Security), versus 6.22 and 5.75 for Group 2. In ITU, the differences persist: 7.00 (Safety) and 6.58 (Security) for Group 1, against 6.63 and 6.67 for Group 2.

In the consolidated TAM, results are split between the groups: Group 1 stands out in Safety (6.97 vs. 6.53) and Security (6.93 vs. 6.49), while Group 2 shows slightly higher scores in STPA (6.64 vs. 6.58) and IoT (6.64 vs. 6.54). The same occurs with TTF, where Group 1 performs better in Safety (6.90 vs. 6.02), but Group 2 leads in Security (6.80 vs. 5.95), STPA (6.20 vs. 6.17), and IoT (6.20 vs. 6.18).

The analysis of responses by declared level of knowledge/experience in the areas of interest revealed only very subtle variations in tool evaluations across the TAM and TTF constructs, regardless of participants' expertise in each area.

Thus, based on the data analyzed, it is not possible to state with statistical validity that certain expert profiles are more or less favorable for using the *SafeSecRETS* tool. The differences observed across declared knowledge levels are small and remain in the upper range of the Likert scale (close to 7), indicating high overall acceptance. In summary, the variations are neither statistically significant nor consistent enough to conclude that one profile is more suitable or that the tool should be recommended to a specific group. Instead, the evaluation highlights broad and homogeneous acceptance, regardless of prior knowledge in the areas considered.

Discussion

The analysis of the influence of participants' knowledge level on the evaluations revealed subtle differences in perception between the conceptual artifacts (*SafeSecIoT Canvas* and *STPA-SafeSecIoT*) and the *SafeSecRETS* tool. The results indicate that, although all artifacts were positively evaluated across constructs, distinct patterns emerged in how Groups 1 and 2 perceived PU, PEOU, and ITU.

For the conceptual artifacts, participants with less declared knowledge in IoT and STPA assigned higher scores in all constructs, both for the *SafeSecIoT Canvas* model and for the *STPA-SafeSecIoT* method. This suggests that these artifacts serve as structured guides that help reduce the complexity of safety and security analyses, particularly for less experienced participants or multidisciplinary teams. Among experts, although evaluations remained highly positive, perceived gains were slightly lower, possibly due to their prior familiarity with the concepts and practices operationalized by the artifacts. Therefore, the

conceptual artifacts prove valuable not only for supporting the safety and security RA process of critical IoT systems, but also as pedagogical and methodological resources, being recognized as useful and easy to use even by more advanced profiles.

In the case of the *SafeSecRETS* tool, construct evaluations were higher and more consistent across Groups 1 and 2 than those for the conceptual artifacts. Mean scores remained high for both groups, highlighting the tool's usefulness and its intuitive, easy-to-navigate interface. ITU scores were very high across all profiles, reaching the maximum of 7.00 among Group 1 participants for Safety, Security, and IoT. These results indicate strong potential for practical adoption, with experts valuing the tool regardless of their experience in IoT or STPA. Participants recognized its structured support for safety and security RE activities, complemented by collaborative features, traceability, partial automation, and overall process assistance.

Overall, the findings suggest that the three evaluated artifacts are highly useful, easy to use, and likely to be adopted, each fulfilling complementary roles in supporting the safety and security RE process. The conceptual artifacts provide a solid methodological foundation, facilitate understanding, and enhance communication within heterogeneous teams, making complex analyses (such as STPA-based evaluations) more accessible. *SafeSecRETS* represents the practical implementation of this process, offering features that support adoption in real critical IoT system development scenarios. This complementarity, along with the consistently positive evaluations, underscores the robustness of the proposal, which provides both methodological guidance and a technological solution capable of assisting analysts with varying levels of experience.

9.3 Qualitative Analysis: Open-ended Questions

In order to assess experts' perceptions of the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts, as well as the *SafeSecRETS* tool, responses to open questions about their strengths, limitations, and possible improvements were analyzed. The responses collected offer a view based on the experiences of different experts on how each artifact contributes to supporting the safety and security RE process in critical IoT systems. Below, we present the main findings organized by topic, highlighting recurring and representative aspects in the participants' responses, in order to highlight both the perceived benefits and the challenges and opportunities for improving the approach. This analysis aims to answer research questions **RQ5 and RQ6** defined in subsection 9.1.4.

9.3.1 Open-ended questions: *SafeSecIoT Canvas* model

The experts highlighted that the *SafeSecIoT Canvas* model assists in the initial identification of critical requirements that often go unnoticed or are not captured during the RE process. Its value for novice professionals was also emphasized, as it serves as a learning resource. These points reinforce that the model has a direct impact on the practical usefulness of the RE process, both in academic and professional contexts.

Visual clarity and information organization were identified as factors that facilitate the application of the artifact. In addition, the canvas format was noted as a differentiator for its adoption, as it combines visual and structural aspects. On the other hand, some limitations reveal points of attention for usability: dependence on the user's prior knowledge, loss of clarity in very large analyses, and terminological ambiguities. These aspects suggest that PEOU is perceived positively but can be improved with guidance, examples, and greater conceptual standardization.

The acceptance of *SafeSecIoT Canvas* is strongly linked to its potential for communication and collaboration, especially for non-specialized audiences. The perception that the artifact serves as an organizational guide for the process and facilitates the integration of different perspectives reinforces the positive trend in its use. Suggestions for improvement (such as including more practical examples, proposing an information catalog, and even considering automated filling) reveal that experts not only recognize the value of the artifact, but also demonstrate interest in its evolution, which reinforces the intention for future use.

Strengths of *SafeSecIoT Canvas*

The main strengths identified by the experts are summarized below, with supporting excerpts from their responses:

- **Initial identification of essential requirements**: as mentioned by one expert, in addition to supporting the elicitation of essential system requirements, *"The artifact provides essential points for identifying the initial requirements for the safety and security of IoT systems, which are often not captured during the RE stage."*
- **Support for beginners and the learning process**: one evaluator pointed out that, *"For novice professionals, [the artifact] is essential for providing 'skills' and 'experience' in important items to be identified."*
- **Organization of activities and analysis guidance**: this was highlighted by one participant *"The organization of the security analyst's activities. [The artifact] serves as a guide for more complex analyses."*
- **Ease of use and communication of results**: one evaluator stated that the model *"Greatly facilitates the entry of information for analysis, encompassing the initial*

items in a single artifact. This makes visualization quick and practical.'' Another expert adds: *"Easy visualization of information, ease of communicating safety and security analysis results to people outside the field and beginners."*

- **Visual organization and understanding**: one expert highlights the *"Good visual organization and easy understanding."*
- **Potential for adoption due to the canvas format**: one expert highlighted the uniqueness of the proposed artifact, due to its format and ease of use, noting that *"The canvas format is what most increases its potential for adoption."*
- **Structured visual approach and multiple perspectives**: the evaluator states that *"The canvas approach combines visuals with structure, enabling preliminary technical analyses from multiple perspectives."* This is complemented by: *"The metamodel provides important components for safety/security and IoT analysis."*

Limitations of *SafeSecIoT Canvas*

Limitations observed by experts:

- **Loss of clarity in very large analyses**: one of the experts observed a limitation related to the space for filling in and viewing information on the canvas. He states that: *"If the analysis is very large, you may lose the advantage of viewing it on a single screen."*
- **Lack of taxonomy for classifying requirements**: another limitation identified was related to the classification of requirements. Evaluator's observation: *"I did not see a taxonomy that allows requirements to be classified by their nature as functional and non-functional."*
- **Duplicate/confusing terminology**: one of the experts made an observation about the term risk, used in two building blocks. He noted that: "*Since there are two components called 'Risks', this may confuse users."*

Suggestions for Improvement of *SafeSecIoT Canvas*

Suggestions for improvement identified by the experts:

- **Adjust the size or scope of the canvas**: due to the number of building blocks in the proposed model *"One suggestion for improvement, perhaps, would be to make it a little smaller (I don't know if that's possible)."*
- **Include practical examples**: another suggestion for improvement was related to examples for guidance on filling out the canvas. The evaluator states that *"It would be interesting to include examples [of filling out the canvas]."*

- **Propose a catalog of supporting information**: another suggestion to support filling out the canvas was to create a catalog. *"Suggestions for improvement, for future work, propose a catalog of information whenever possible."*
- **Automate filling in based on previous analyses**: another suggestion related to filling in. The expert states that: *"It may be possible to automate the filling in of the canvas, based on a previous analysis."*

9.3.2 Open-ended questions: *STPA-SafeSecIoT* method

The method was valued for allowing a correct and standardized specification of safety and security requirements, highlighting its consistency. Another useful aspect was its ability to integrate safety and security, including intentional and unintentional threats. In addition, it was also highlighted that the method maintains the essence of STPA, without deviating from its original objective. These points reinforce the PU in contexts of systematic, consistent, and integrated analysis of critical requirements.

Despite recognizing the value and PU of *STPA-SafeSecIoT*, the learning curve was pointed out as an obstacle. Another critical point was the difficulty in understanding the analysis flow and the absence of templates (as in the case of canvas) to support the application of the method's steps. It was also noted that some elements, such as loss scenarios, need to be more concrete. These comments suggest that, although the method is perceived as rigorous and useful, its PEOU depends on auxiliary resources (documents, templates, examples).

The ITU use the method appears to be associated with the perception that the proposed method is an integrated and innovative artifact that expands STPA to address both safety and security. Suggestions for improvement reveal a predisposition to use the method, provided there is practical support for its implementation, such as a tool, for example. This shows that experts see value in the method and would like to see it strengthened with complementary resources that reduce the complexity of application, which reinforces their intention to use it in the future.

Strengths of *STPA-SafeSecIoT*

- **Correct and standardized specification of requirements**: an expert states that the method *"Shows how safe and sec requirements are specified in a correct and standardized manner."*
- **Consistency and completeness of the method**: the evaluator highlights how *"Strengths: complete and consistent."*
- **Integration of safety and security**: the expert highlights that *"The method addresses threats of an intentional and unintentional nature. The division into Unsafe*

and Unsecure actions allows the analysis to be divided between safety and security.'' This is complemented by another assessment: *"An integrated method for safety and security, and use of the security process to analyze safety, something like the STPA-Sec method set out to do."*

- **Clarity in the integration between safety and security**: one of the evaluators highlights that the method *"clearly shows the connection/division between safety and security elements."*
- **Detailing the steps of STPA, facilitating understanding of the technique**: in this evaluation, the expert highlights that the proposed method *"Details the steps of STPA, which is a technique that is difficult to understand, especially for beginners."*
- **Alignment with the original objective of STPA**: an important point noted is that the proposed method *"Adequately addresses the original objective of the STPA method, without deviating from the focus."*
- **Visual organization and understanding**: one of the evaluators states *"Strengths: good visual organization and easy to understand."* Even though it is a method with several steps and tasks, it was considered easy to use by the expert.

Limitations of *STPA-SafeSecIoT*

- **Dependence on prior knowledge and steep learning curve**: in contrast to an expert who already had prior knowledge of STPA, another evaluator states that the use of the method *"Depends on the user's prior knowledge, the learning curve can be a limitation.''*
- **Dependence on the canvas for performing the analysis**: it was also noted that *"This second part can be a little more 'complicated', as it depends entirely on the SafeSec[IoT] Canvas.''* This assessment highlights the integration and consequent dependence between the artifacts.
- **Lack of support templates for recording the analysis**: one expert noted *"As a limitation, I thought about not having an artifact [template] to be filled out at the moment, so the analyst needs to create the document himself.''*
- **Difficulty understanding the analysis flow**: one of the evaluators reported difficulty understanding the method. *"I confess that I was a little lost in relation to how the analysis flow itself would be just by looking at the method, but I managed to understand it with the explanation.''*
- **Lack of clarity in the application of loss scenarios**: one evaluator stated that *"In the Task: Identify Loss Scenarios, it was not possible to visualize the concrete application, as I could not find the artifacts instantiated in this task.''*
- **Restriction regarding focus on customer requirements**: one evaluation participant noted that the method also needs to consider the requirements inherent to the

system. *"Do not limit yourself to focusing on customer requirements.''*

Suggestions for Improvement of *STPA-SafeSecIoT*

- **Inclusion of examples and greater detail in the scenarios**: one expert states that *"It would be interesting to include examples.''* and adds that *"The scenarios section needs to be better explained."*
- **Automated tool or support for analysis**: this evaluator notes how *"Suggestions for improvement, for future work, a tool to support the analysis and specification process."*
- **Creation of templates or pre-structured documents**: the evaluator noted that it would be interesting to include templates for the steps of the proposed method. *"A suggestion for improvement would be to create a document or environment pre-filled with titles, subtitles, and tables [for the data] to be added."*
- **Improvement in the analysis flow**: the expert reported *"difficulty in understanding the analysis flow. Suggestions: improve the flow."*
- **Clarity in differentiating between original STPA activities and extensions**: the evaluator suggested *"Making it clearer to those familiar with STPA which activities are new (not in the original), which have been adjusted, and which are activities from the original method."* In this case, to maintain the structure of the original method, the original activities were adopted and adjustments were made to the tasks, expanding the scope to address security in addition to safety.

9.3.3 Open-ended questions: integration between artifacts

The experts were also asked about their perceptions regarding the integration between *SafeSecIoT Canvas* and *STPA-SafeSecIoT*. Below are the responses organized by topic:

- **Recognition of the value of integration**: experts emphasized that integration between the two artifacts *"makes perfect sense"* and can bring *"great potential benefits for the development of critical IoT systems"*. There were also positive perceptions regarding the value of the canvas as an initial step that underpins and guides the more detailed analysis of *STPA-SafeSecIoT*: *"I found it very interesting to have this first step worked out in the* SafeSecIoT Canvas*, since the basis of the analysis is extremely important and guides the entire analysis"*. It was added that *"The canvas itself already adds value, helping to initiate a preliminary analysis. The effectiveness of integration with the Analysis/Specification artifact will depend on tool support, but the process makes clear what the inputs/outputs are between the two artifacts."*

- **Need for greater clarity in integration**: despite recognizing its relevance, some experts indicated that the integration is not yet sufficiently clear. There were suggestions to include practical examples and traceability mechanisms to better explain how the results of one artifact feed into the other: *"The integration is not so clear. Examples could help make the integration more visible. How to show tracking?"*. Another evaluator made the counterpoint, but reinforced these needs: *"I saw that it seems to be perfectly traceable, but I confess that I would need to see it in more detail to understand how traceability would work."*
- **Requirement for prior experience and knowledge**: integration was perceived as positive, but still geared toward professionals with more experience in safety and security, which may represent a barrier for beginners: *"Good, but it seems to be aimed at experienced professionals."*
- **Challenges of consistency between safety and security**: experts drew attention to the need to deal with conflicts between requirements arising from the two perspectives (safety and security), which reinforces the importance of methodological mechanisms and support tools to maintain consistency: *"The division between safety and security requires greater organization of the database and integration between the requirements generated by the two analyses, since the requirements arising from both may conflict with each other. Thus, the method needs to consider the resolution of conflicting requirements."*
- **Suggestions for future development**: improvements were suggested, such as adding new automated steps to the canvas to facilitate analysis and reduce the analyst's effort, as well as adopting bidirectional approaches (backward and forward) to reinforce traceability between artifacts: *"I believe that as future work, you can add other steps to the canvas and generate a more automatic analysis. Only more complex steps would be left to the analyst. Because facilitating the use of STPA can be very interesting. Also, showing different ways to use its techniques, because from my point of view, it serves in various ways: for a person who already has a ready analysis and wants to refine it, for projects that separate the safety/security part from the part that does not have these characteristics, [etc.]."* Furthermore: *"It seems to me that integration makes perfect sense, and if possible, following the backward and forward idea."*

It was found that experts recognize the integration between the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* as promising, useful, and capable of adding value to the RE process in critical IoT, especially by providing an initial planning flow that supports detailed analysis. However, they point out the need for greater clarity, practical examples, traceability mechanisms, and tool support, in addition to attention to resolving conflicts between safety and security.

9.3.4 Open-ended questions: the *SafeSecRETS* tool

In addition to questions based on the TAM and TTF models, the questionnaire included an open-ended question asking experts to indicate the strengths, limitations, and suggestions for improvement for the *SafeSecRETS* tool. The experts' responses demonstrated a largely positive perception, highlighting its role as a facilitator of the safety and security RE process in critical IoT systems, facilitating the integration and practical use of the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts. Below we discuss the responses collected in the survey.

Strengths of *SafeSecRETS*

The experts pointed out the following strengths: i) the intuitive and easy-to-learn interface; ii) the fluidity of use; and iii) the structured support for identifying system elements: assets, system components, hazards (safety-oriented), and threats (security-oriented). The *SafeSecRETS* tool was recognized for its usability and intuitive interface, making the process of engineering safety and security requirements clearer, faster, and more practical. Below we highlight the strengths and present the comments of the experts, highlighted in italics:

- **Ease of use and intuitiveness**: the responses highlight the ease and intuitiveness of the tool, with one expert stating that: *"the tool is very coherent and intuitive"*; which is complemented by another evaluator (without experience in STPA): *"the negative perception of the STPA artifact disappeared when using it (in fact, all that complexity is not necessary for something so straightforward)."* Another expert reinforces that *"the use of the tool [is] very fluid, with well-defined steps."* Usability is highlighted: *"the strong point is usability."* The expert adds: *"I believe it may be interesting to define or identify how to use the tool to elicit, i.e., during an interview, workshop, or focus group."*
- **Support for identifying elements relevant to the system**: the tool is effective in *"supporting the identification of system assets and components as well as safety-related hazards and security threats."*
- **User interface (UI)**: The canvas UI is highlighted for *"facilitating the process"*, and the *"implementation of the STPA support UI facilitates the filling in of information."* The interface and the generation of the control structure are also highlighted as strengths of the tool.
- **Collaboration**: a key factor is the tool's ability to be collaborative, with one expert stating that *"an entire team can use the same project."*

- **Potential for AI use**: the integration of artificial intelligence is seen as a differentiator, as *"based on the information included in the project, AI can assist in generating the requirements."*
- **Traceability**: The tool makes traceability more concrete, as one of the evaluators states: *"The traceability part became concrete for me. In the video of the approach, I couldn't visualize what the traceability and instance of the artifacts would look like. These points only became clear with the tool."*

The combination of clarity, structure, and practicality was frequently associated with a reduction in the perceived complexity of applying the *STPA-SafeSecIoT* method, which some experts had initially considered excessively detailed. Another recurring highlight was the tool's collaborative functionality, enabling multiple team members to work on the same project simultaneously, along with its direct support for element traceability and systematic risk analysis.

In addition, some experts identified specific features as distinctive strengths of the *SafeSecRETS* tool, such as: i) the integration of artificial intelligence (AI) to assist in generating constraints and requirements from project information, and ii) the linkage between the canvas interface and the control structures derived from STPA, considered valuable for ensuring artifact consistency. These positive perceptions are consistent with the high scores obtained in the quantitative evaluation, particularly in the PEOU construct of the TAM model, reinforcing that the tool is regarded as both efficient and user-friendly, even for complex tasks in the context of critical systems engineering.

Limitations identified of *SafeSecRETS*

Despite the predominance of favorable assessments, limitations and concrete opportunities for improvement were identified. Among the technical aspects, the following stood out: i) the absence of automated reports and data export (e.g., in JSON formats); ii) limited traceability visualization, currently restricted to direct connections without exploring transitive relationships; iii) and performance issues, with response times perceived as high in certain actions. Another limitation pointed out was the absence of automated requirement classification mechanisms (such as the distinction between functional and non-functional requirements), which was highlighted as important to complement the complete requirements specification. Terminological revisions were also suggested, such as replacing *System Risks* with *System Hazards*, and the explicit separation between *hazards*, *damage scenarios*, and *security threats*, seeking greater conceptual precision.

Below are the limitations, with highlights from the experts' comments:

- **Report generation and/or data export**: the most frequently mentioned limitation is the lack of functionality to *"generate the report"* or *"export the data in report*

or json format." However, this is considered a minor issue compared to the overall quality: *"I believe the only limitation is generating the report, but it will be resolved and is certainly a very minor point compared to the overall quality."*

- **Information in the interface**: it is suggested that *"[...] the interface should have more information on control actions and components.''* In addition, it was highlighted that *"the UI only shows direct (adjacent) traceability between elements, it could show transitive traceability to more important elements (Hazards/Threats)."*

Suggestions for improvement of *SafeSecRETS*

The suggestions for improvement presented provide guidance for future versions of the tool, aiming to enhance its functionality, usability, and integration with other tools. Below are the suggestions for improvement, highlighting the comments made by the experts:

- **Separation of concepts**: one expert suggested *"separating the specification of hazards, damage scenarios, and security threats*".
- **Generation and classification of requirements**: the lack of *"generation of requirements to compose the complete requirements specification and a classification of SS requirements into functional and non-functional"* in the tool was highlighted.
- **Terminology**: suggestion to *"replace the term System Risks with System Hazards."*
- **Inclusion of *stakeholders***: an expert suggested *"bringing stakeholders into the analysis of Losses."*
- **Navigability and performance**: an expert identified that there is a perception of *"a slightly long response time for certain actions.''* It was suggested to *"apply usability heuristics to precisely identify the points where usability becomes difficult."*
- **Detailed information in the interface**: to improve the interface with more information (*control actions*, components), it is suggested to *"create expandable menus to get around the problem"* of visual pollution.
- **Integration with other tools**: an expert asked *"How would it be possible to integrate this tool with other tools already in use in the software development cycle¿'* Integration is seen as *"an essential factor for the adoption of the tool in practice."*
- **Clarification of benefits and gains**: there is a need to *"establish more explicit criteria that indicate the benefits of using the approach and support tool instead of using the techniques and tools already adopted in the critical IoT systems industry.''* It is essential to articulate *"What gains does your work bring in relation to what already exists?"*

- **Guidance for elicitation**: it is suggested that it would be *"interesting to define or identify how to use the tool for elicitation, i.e., during an interview, workshop, focus group"*, improving the guidelines for using the tool for the elicitation activity.

The experts recommended integrating *SafeSecRETS* with tools already used in the industry, with a view to reducing barriers to adoption in real environments, and applying usability heuristics to identify specific navigation points that could be improved. Other recommendations included: i) creating expandable menus to present additional information without compromising the clarity of the interface, and ii) establishing explicit criteria that demonstrate the benefits of using the approach and tool in relation to established practices, especially to justify its adoption in industrial projects.

Discussion and summary of the analysis

The *SafeSecRETS* tool is considered a *"great result of the work"* that *"exceeds initial expectations."* Its features, such as intuitiveness, PEOU, and support for identifying system elements, are important pillars, according to experts. The proposed improvements, particularly in data export, interface detailing, performance, and, crucially, integration and communication of benefits, are important for its adoption and continued success in industrial environments, especially in critical IoT systems. Meanwhile, the collaborative capacity and potential for AI use are promising differentiators for the tool's future.

The additional observations provided by the experts reinforce the quantitative results, especially in the TAM and TTF model constructs. The high PU was emphasized in comments that highlighted the organization of the interface, systematic traceability, and structured storage of elements, pointing to concrete gains in conducting safety and security requirements engineering. The PEOU was corroborated by the evaluators' recognition of the tool's visual clarity, step-by-step structure, and intuitive design, facilitating its adoption even by users who intend to apply the original STPA. The ITU was evidenced by interest in its practical application, including suggestions for empirical studies in educational environments and adoption by teams that already use STPA.

From the point of view of task-technology fit (TTF), the experts emphasized the tool's adherence to analysis tasks, highlighting features that directly support traceability, data storage, and integration with essential STPA components, although they suggested the inclusion of additional elements, such as external communication, which may represent relevant vectors of vulnerability. Additional recommendations, such as the availability of a demonstration video and support for the implementation of other STPA extensions, reinforce the alignment between task requirements and the resources offered, indicating that the *SafeSecRETS* tool has a high degree of suitability and potential for practical adoption by safety and security requirements analysts.

In summary, the qualitative analysis reinforces the quantitative results by highlighting an overall positive perception of *SafeSecRETS*, with high PU, PEOU, and ITU, as captured by the TAM constructs. Criticisms and suggestions, while relevant, focus on incremental adjustments that do not compromise the core functionality of the tool. This combination of high acceptance, intuitive usability, and a clear set of potential improvements suggests that *SafeSecRETS* is well positioned to effectively support the safety and security RE process, with good prospects for adoption and continuous evolution.

9.4 Qualitative Analysis: Meeting Transcripts from the Evaluation Sessions

In addition to open-ended questions answered by the experts, we also explored the contributions made during the evaluation meetings through qualitative analysis. The purpose of the evaluation meetings was not to interview the experts. For this reason, no specific questionnaire or set of questions was defined or applied in order to structure an interview. However, at the beginning of each evaluation, we informed participants that they could make comments and contributions on the artifacts evaluated throughout the process. The evaluation meetings followed the same script and lasted between 70 and 120 minutes, depending on the pace, comments, and contributions made spontaneously by each participant. This qualitative analysis considered the transcripts of the 11 evaluations carried out in order to analyze the detailed perception of all experts in relation to the artifacts evaluated.

The preparation of data for the coding stages involved the transcription of the 11 meetings conducted. Each meeting was transcribed into a text file that included the identification of the statements made by each participant (researcher and participating expert), ensuring traceability for subsequent analysis and coding. The analysis and coding process, based on Grounded Theory (GT), was organized into two stages:

- **Open coding stage**: in this stage, the transcripts were read in detail in search of relevant contributions made by the experts. Each relevant quote was assigned a code, which made it possible to name the points that the experts were interested in or considered relevant during the evaluation.
- **Axial coding stage**: the objective was to identify the categories into which the different codes fell and to group them together. In addition, we sought to define the relationships between the codes, according to the data extracted from the transcripts.

The essence of the GT method lies in the fact that theory must emerge from the data, that is, derive directly from the information collected and analyzed system-

atically [Conte et al., 2009]. However, although the central purpose of this methodology is to construct theories, researchers may choose to use only some of its procedures in the data analysis process, according to the specific objectives of the research [Corbin and Strauss, 2014]. In the present study, as only one cycle of data collection was conducted, the phases of open coding and axial coding were applied to the transcripts obtained from the 11 meetings held in the context of this evaluation.

The analysis and coding work was carried out with the aid of the ATLAS.ti tool[2], version 9, which is one of the best known and most widely used software programs for data analysis in qualitative research, due to its powerful features, including different types of coding [Hwang, 2008]. It is important to note that the tool assists the researcher in the process of organizing data analysis, but does not perform any type of autonomous analysis. All inferences and categorizations must be made by the researcher, supported by their theoretical basis, promoting an interface between human expertise and software data processing.

The open coding in the tool aimed to identify relevant information that could help answer the questions defined for the evaluation. For this purpose, codes were created that assign a meaning to each highlighted quote, as shown in Figure 9.8. In this process, it was also found that there were redundant codes (with the same meaning). To refine and improve the codes, they were reviewed by two researchers external to the work, with the aim of eliminating redundant codes and identifying and improving ambiguous or imprecise codes.

After open coding and peer review of the codes, axial coding was performed, aiming to create groups of related codes representing categories of information. As this was an evaluation of artifacts in support of a safety and security RE process, the main categories created based on the experts' contributions were defined as: strengths, limitations, and suggestions for improvement. Subsequently, subcategories were defined according to the quality characteristics and other aspects observed.

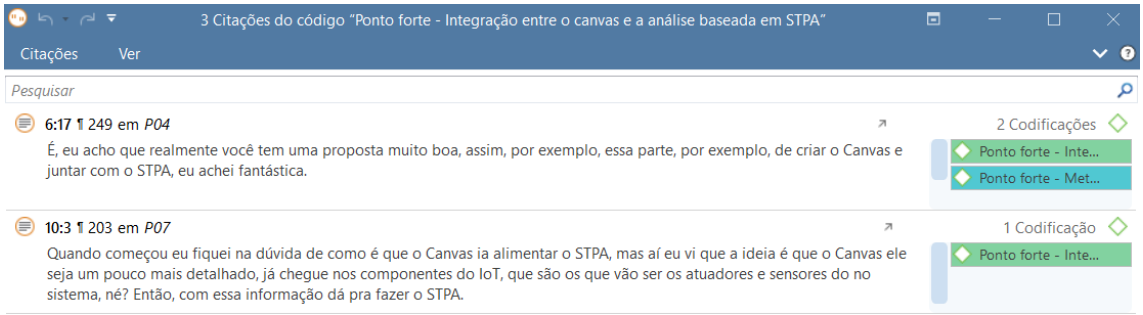


Figure 9.8: Example of open approval assigning codes to two obligations of different participants.

[2] https://atlasti.com/pt

For better visual representation and analysis in line with the defined research objectives and questions, networks were generated for each dimension analyzed during axial coding: positive points, suggestions for improvement, and limitations of the work. In this way, the relationships between the codes were examined, allowing for analysis based on the relationships between what was observed by the experts.

9.4.1 Strengths of the proposal

Axial coding of GT [Corbin and Strauss, 2014] is the stage in which we seek relationships between the initial (open) codes, organizing them into more abstract categories that help to better understand the collected data. The reasoning follows a logic of questions that aim to relate the data:

- What do the codes have in common? Do they have similar content or focus?
- What differentiates them from other groups of codes? Observed characteristics: contrast, scope, object of analysis.
- What do they describe: a characteristic of the artifact, a perceived benefit, usability, a technical issue, a methodological aspect, etc.?

In this case, the 32 codes previously classified as strengths/positives were analyzed to identify categories. The logic used to generate categories was: i) identify convergences of meaning; ii) separate process and product aspects; iii) look at purpose/impact; and iv) seek levels of abstraction that can generate groups. In axial coding, general codes become broader categories, and specific codes can become subcategories or examples that embody the category.

Figure 9.9 shows the network created through axial coding in order to answer **QP5 – What are the strengths of the safety and security RE process for critical IoT systems based on the proposed artifacts?**

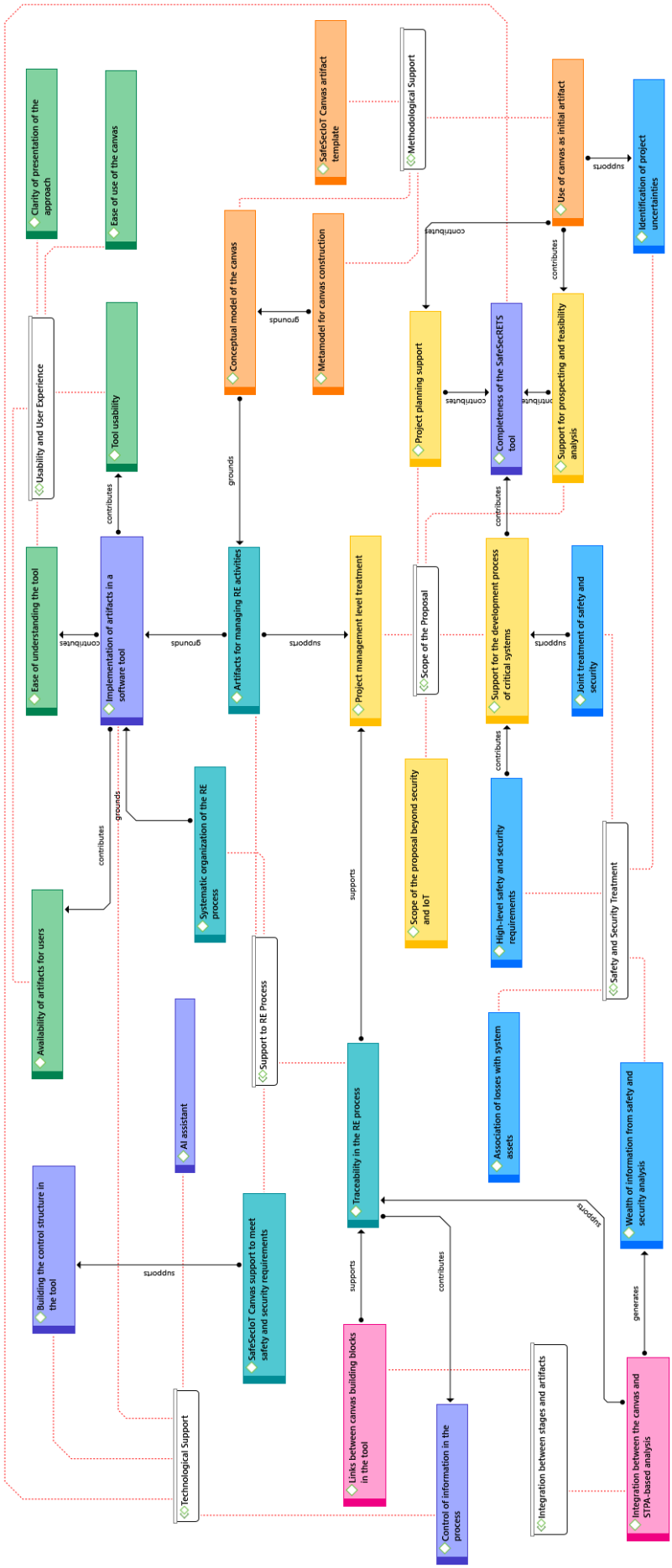


Figure 9.9: Positive aspects (strengths) of the proposed approach, according to experts.

According to this methodology, the identified groups were:

- **Proposal Scope**: brings together points that highlight the breadth of the work, the coverage of important aspects, and the ability to support different tasks related to RE;
- **RE Process Support**: includes points that highlight direct support for formal RE activities, from elicitation to requirements management.
- **Safety and Security Treatment**: includes points that highlight direct support for formal RE activities, from elicitation to requirements management.
- **Methodological Support**: refers to aspects related to the conceptual and formal design of the proposal, the canvas as an artifact, its internal organization, and modeling rules based on the metamodel.
- **Technological Support**: brings together points related to software materialization, technological support, and additional control mechanisms.
- **Usability and User Experience**: groups the factors that facilitate practical use, understanding, and interaction with conceptual artifacts and the tool.

Proposal Scope

Regarding the scope of the proposal, the main point highlighted by the experts was the *support for project planning* and how it demonstrates support for the stages of elicitation, analysis, and specification of safety and security:

- P1 - *"[...] I liked the What, Who, Why, which is the 5W, cool. Very cool! That helps at the beginning... Before the person [analyst] starts the two cycles, right, the safety-security life-cycle, this gives basic information so that the person has an idea and can already conduct the hazard analysis and threat analysis activities."*
- P6 - *"[The proposal] helps [...] to pick up issues that we may not see at the beginning [of the STPA analysis]."*
- P9 - *"I think what you did has a lot of advantages because people, they... I think it's harder to understand what STPA [can analyze] with this basis. If we do this basis well, you don't lose any details for the next steps, and then you create the entire [analysis structure] for the person."*

Still in relation to project planning, two experts also highlighted the *scope of the proposal based on the metamodel and its instances*, assessing that it could be used more broadly for RE in other types of projects, even if they did not involve safety and security, or even beyond the domain of IoT:

- P3 - *[...] gave me the impression that your proposal is more comprehensive than just safety and security, you know? So, I see a very interesting structure for building a*

canvas for projects. And in this project canvas, when identifying safety and security, I have solid support to meet my requirements, which is not trivial. It's not trivial."

- P9 - *"And then you could have that [metamodel instance] for any [domain]. Instead of IoT, you could have it for any critical system [...] Your contribution was much greater, wasn't it? Your contribution was much greater than IoT, right? [...] But I think it's the greatest contribution because of its generality, not that the other [artifacts] aren't... The other [artifacts] are very good too, but this one shows how much you've created a general solution, right?"*

One of the experts also highlighted *"the support for the project's feasibility analysis"*, particularly with regard to identifying safety and security issues that could hinder the system's development:

- P3 - *"It's the starting point, that is, we are doing systemic prospecting and I want to build an initial vision to see if this [project] is even feasible, because I may have so many requirements... so many security and possibly safety issues that I'll say, man, give up on this system, it's not going to work this way.*

Finally, one of the experts highlighted the *coverage of the project's management level* (based on the aggregation of the canvas-based artifact) with the discovery of information related to the scope, and the top-down approach that supports the STPA-based analysis performed later:

- P4 - *"When you take the part of the canvas you created before, where you put that visual organization, where you separate the rectangles and such. I thought that was really good, I thought that was really good. [...] So, in that respect, I think you have a path where, when you start with the canvas, which is something that gives you a level of management, a more managerial level, a level of completeness, you see a lot of things integrated there, I think it takes away a little from that idea that STPA is very detail-oriented. Because you start from a slightly higher level view and then you go to something a little more down to earth to get into the details, right? So I think that's one of the very strong advantages of your work. Very strong."*

In the analysis process, we also considered the relationship between the codes of each group and the others. For example, support for the development of critical systems, project planning, and prospecting and feasibility analysis contribute to the completeness of the *SafeSecRETS* tool developed (Technological Support). In addition, the use of canvas as an initial artifact (Methodological Support) directly contributes to the support for project planning and feasibility analysis highlighted in this group. We also note that the joint treatment and specification of safety and security (Safety and Security Treatment) are also related to the support for the critical systems development process.

RE Process Support

In this second group of codes, five experts highlighted in their statements the *support for the safety and security RE process and the systematic organization* carried out in stages through the artifacts. Some excerpts that exemplify this perception:

- P1 - *"You can organize project information, right? That's cool. I think that's the goal, right? Of this methodology, which consists of the metamodel, the process, and the tool, right¿'*
- P3 - *"Because then things start to make sense, because in fact what seems complex is not complex, it's just very well organized. Oh, now you do this, that's it. Now you do that, that's it.''*
- P8 - *"So, what I realized is that [...] you took great care to ensure that everything was consistent, that is, it seemed to me that your work was thought out in a systematic way, right¿'*

Another positive point widely observed and cited by participants was the traceability between artifacts and the information generated, explicitly highlighted by five experts. Some excerpts from the comments on this point:

- P3 - *"So you show the trail, if you show that trail, then you will show the evolution of that business until you reach the requirement. [...] Because in the end, you see, in the end you have such a wealth of information in that requirement and in the trail that is what you want to achieve.''*
- P8 - *"[...] so in terms of security, I see that you really tried to be very thorough and consistent throughout the entire process and that you really tried to present artifacts that would allow me to manage what will be the main activities of requirements engineering, and so at least in this way, overall, I think this is very positive, especially the issue of traceability.''*
- P9 - *[...] but it ended up being very married. This issue of traceability is something that is impressively interesting.''*
- P11 - *"[...] the way you link [the information], your work is very good.''*.

The artifacts for managing RE activities are based on the conceptual canvas model (Methodological Support) and support the treatment of the project management level (Proposal Scope), highlighted as an important contribution of the approach. In addition, the systematic organization of the RE process supports the implementation of artifacts in a software tool (Technological Support). Finally, we highlight that the support of the *SafeSecIoT Canvas* to meet safety and security requirements supports STPA-based analysis, including the construction of the system control structure.

Safety and Security Treatment

One of the experts highlighted the challenge of treating safety and security together:

- P8 - *"Because I think that your work, in terms of safety and security, is really something extraordinary, you see? You have, that is, your work is phenomenal and is indeed very important, you are here doing a tremendous job."*

Experts highlighted the wealth of information produced during the process to support safety and security analysis:

- P1 - *"Components, actions, right? The hardware, the software to control glucose, the embedded CGM and the pump... server, application for remote monitoring, and the actions are reading glucose, checking insulin needs, and administering insulin. Three actions: patient glucose, amount of insulin, insulin dose, and connectivity. Cool. Here you can already get project information [to analyze safety and security]."*
- P3 - *"Because in the end, you see, in the end you have such a wealth of information in that requirement and in the trail that you want to follow."*

- P9 - *"That requirement... you call it a requirement because I could already end there, because we have a certain problem with our scenario, not just us, everyone, right? [...] But so, it's a methodology that seems to me that you could stop if you wanted something high level there, right? You didn't even need to go to the scenarios. [...] Oh, I thought that was great.*

Finally, one participating expert highlighted the importance of identifying system assets before initiating the safety and security analysis with the definition of losses. In this regard, we emphasize that addressing safety and security jointly and specifying their respective requirements provides direct support to the critical systems development process (Proposal Scope). Another relevant point observed is that the integration between the canvas and the STPA-based analysis (Integration between Stages and Artifacts) enriches the safety and security analysis by generating a broader and more detailed body of information.

Methodological Support

The experts highlighted the metamodel and conceptual model developed as important for the methodological support of the approach. Regarding the *MM4Canvas* metamodel, we highlight:

- P3 - *"So, I see a very interesting structure for building a canvas for projects. And in this project canvas, when identifying safety and security, I have solid support to meet my requirements, which is not trivial. It is not trivial."*

The *SafeSecIoT Canvas* was highlighted as an important conceptual model in the context of the proposal:

- P3 - *"[...] for example, that Who, What, Why model. Man, that's a fantastic conceptual model [...] for establishing formalism, right?"*
- P7 - *"Even BMC's original [canvas model] doesn't have a metamodel. Because it's not their area, right? Here, every time we see a diagram, we look at it and say, what is the metamodel of this diagram?"*

In this regard, the use of a canvas as an initial artifact for the RE process was also highlighted, supporting STPA-based analysis:

- P9 - *"[...] I'll have the canvas as a starting point, right? [...] I found it very interesting because agile people, especially, [...] do very well with canvas. I confess that I don't know much about it, but I find it very interesting as an initial artifact. Then, based on that, you [...] do the STPA part, right?*

In this context, the metamodel for canvas construction supports the conceptual model developed, the *SafeSecIoT Canvas*, which is one of the artifacts that support the management of RE activities, in this case project planning and elicitation. The use of the canvas as an initial artifact contributes to project planning and feasibility analysis (Proposal Scope) and supports the identification of project uncertainties (Safety and Security Treatment).

Technological Support

Regarding technological support, eight participants highlighted the importance of implementing the artifacts in a software tool. Below are some excerpts that exemplify this perception among the experts:

- P1 - *"Well, it was really hard work, and it added value to your doctoral thesis, to the process, and to the metamodel you created. You showed that things work, right? Cool."*
- P4 - *"I think the tool provides more... including contributions [to the work]."*
- P6 - *"[...] looking at it this way is much better, right? In the sense that it's more... usable than, for example, the artifact itself. [...] I like it. The tool is much better in this case. [...] A tool that helps us be more efficient and effective in the context of specification, validation, and everything else is better."*

- P9 - *"[...] I think you're going to, how do you say it? You're going to delight everyone with this tool."*
- P11 - *"[...] I found the tool very interesting, especially this part of generating control structure here."*

In addition, another positive point highlighted by the experts was the development of an AI assistant to support some parts of the analysis and specification method in the tool:

- P5 - *"Agora me explica como é que você fez para implementar [o assistente de] AI. Eu vi que você usou Fireworks. Como é que foi isso? [...] ficou muito legal. [...] Uma IA vai ser mais eficiente para poder simplesmente listar tudo que tá pegando, né? Talvez para as UCAs ali, né? Pras unsafe control actions e outra outra IA só pro reasoning."*

The completeness of the tool was also highlighted by the experts:

- P1 - *"[...] your tool is very comprehensive, right¿'*
- P6 - *"A very comprehensive tool. I really liked it. [...] It's really good, very comprehensive, actually, right? I think you can already share it and do several experiments with it. TAM evaluation."*

Control of the information generated in the process was also highlighted:

- P1 - *"Yes, you can model components, component actions, manipulated data, and how components connect to each other [...] so you can define component instances and how they connect, right."*
- P3 - *"This grain [level of granularity] that's here... A1, A2, A3. I mean, that's because when you enter the tool, you enter item by item, right? And it has this control, right?"*

Some relationships can be highlighted in this context. The implementation of artifacts in a software tool is based on the systematic organization of the RE process (RE Process Support) and contributes to the availability of artifacts to users (Usability and User Experience). Also noteworthy is the support of *SafeSecIoT Canvas* for the safety and security analysis process (RE Process Support) for building the control structure in the tool. In addition, the traceability defined between conceptual artifacts (RE Process Support) contributes to the control of process information.

Usability and User Experience

The usability of the *SafeSecRETS* tool was widely highlighted by the participating experts, including as a facilitating element for the application of the artifacts and the proposal:

- P1 - *"This view of the component here [canvas] is really cool. This [part] here already gives you an idea of the project."*
- P3 - *"[...] actually, now [with the tool] I understand it better... I have an initial view of the canvas as a whole, right? And I have the tool, depending on where I click on the canvas, it takes me to where I can detail that group of information."*
- P4 - *"So, I think it's really cool like this. I found it easy to use and everything, making all the associations, right?"*
- P7 - *"You already have a little product there, because it was the first time I saw these... at least this type of interface here to work on a canvas, where you put a theme here and add these little spaces here. I thought it was really cool. I really like using canvas, but one of the difficulties with canvas is showing this whole screen, right? Because you have a lot of different components here, and yet you managed to fit everything in here."*
- P8 - *"[The tool] has a very user-friendly interface."*
- P9 - *"So, besides being aesthetically pleasing, it turned out really well."*

The same perception applies to the ease of understanding the tool:

- P1 - *"Your tool is very intuitive.''*
- P4 - *"Yes, I thought the tool was very good, I thought the tool was very good. It's well tied in with your process... well described, I think it has very good potential for use. I really liked it, it's very intuitive. Right? So, I think it's really cool. I found it easy to use and everything, making all the associations, right?"*
- P7 - *"The tool is easy to understand. If you know which method it supports, you can use the tool, right?"*

The clarity of the supporting material and the video presentation of the approach was also highlighted:

- P1 - *"I understood [the proposal] well. The video was very educational."*
- P3 - *"[...] my comment on the video is that it is very well done."*
- P9 - *"I also found the video very informative."*

Based on this analysis, it can be observed that the implementation of artifacts in a tool (Technological Support) directly contributes to the ease of understanding the proposal, usability, and also to the availability of artifacts to users, in the form of web software.

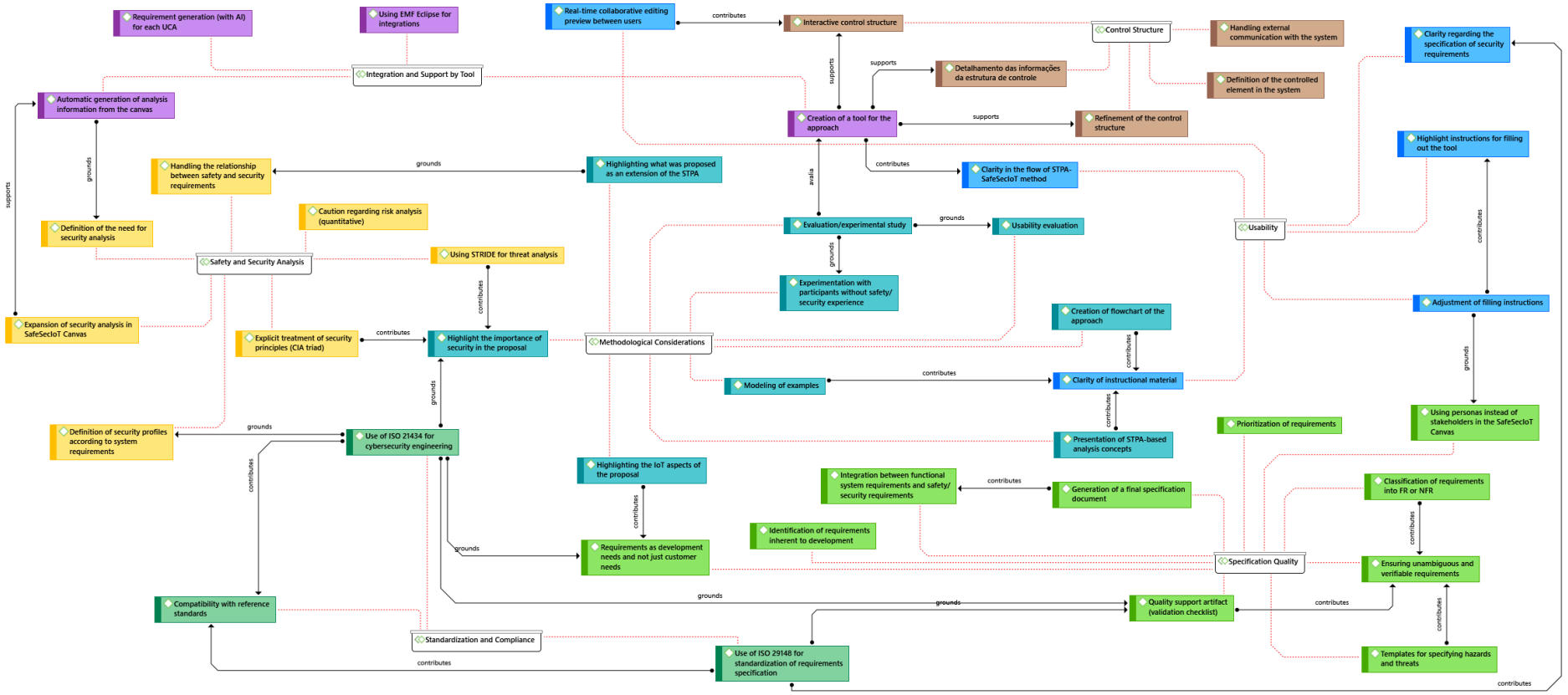


Figure 9.10: Suggested improvements to the proposed approach, according to experts.

9.4.2 Suggested improvements for the proposal

Figure 9.10 shows the network created through axial encoding with the aim of answering **QP6 – How could the safety and security RE process for critical IoT systems be improved?**

Safety and security analysis

The first point in this analysis was the treatment of the relationship between safety and security, highlighted by three participants:

- P1 - *"[...] then you could take it up in a future project, right, and have a single screen with safety and security interplay [...]"*
- P5 - *"So you will need software to manage these two analyses, to make sure you don't end up with conflicting requirements. You may have a safety requirement and a security requirement that conflict with each other, and then you need a new process to help you decide which one will prevail [...]"*
- P7 - *"[...] any tool or process that models [safety and security] separately [...] can be used in this context where safety contaminates security. If this is the case, you generate a version of the security constraint, its safety version [replica], and you will have security treatment for that, but you are also already tracking with safety, so you don't leave the system without this traceability [...]"*

Another point highlighted was the possible use of STRIDE for threat analysis:

- P1 - *"[...] when [analyzing] security, you could also include STRIDE, right? STRIDE is a threat analysis method [...]"*

One participant recommended explicit treatment of security principles (CIA triad):

- P8 - *"[...] in terms of RE, when we think about security, we always tend to think about, and even security experts themselves talk about [...] what we might call the three components of security, which are CIA, right? Confidentiality, integrity, availability... [...] So, my question is, in terms of your proposal, how do you deal with this? Do you treat them differently or have you decided to treat them all the same way?"*

One of the participants suggested improvements in terms of expanding the safety and security analysis based on the *SafeSecIoT Canvas*:

- P1 - *"[...] OK, but you could, because this is free form, right? But you could update it, leave the templates, right? And then do it separately, right? First for safety risk, then for security risk, understand¿'*

Standardization and Compliance

The use of certain standards was suggested to improve quality and maintain compliance with specifications:

- P4 - *"Because once your thesis is ready and you have well-specified safety and security requirements as a good starting point, you have... Then you will have, for example, let's say, a possibility for future work that would be to see the compatibility of these requirements that are extracted using your proposal, so that they are compatible with some standards in specific domains, because several of them deal with very similar characteristics that would easily allow you to include this."*
- P4 - *"You have to set up a nomenclature that standardizes these requirements. So, for example, normally in requirements engineering we have standards that guide the structure of a requirement. One of them, which I can send you later if you want, is ISO 29148, which replaced IEEE 830, which is a classic standard for describing and stating requirements."*
- P1 - *"Perhaps here you could include a recommendation, not for now, but for later, you could take a look at ISO 21434, which has other information, right?"*

Specification Quality

In this group, we highlight the participants' recommendation to generate a final specification document, including the general system requirements obtained from the canvas:

- P3 - *"Because here we are specifically dealing with safety and security requirements. But what about my general requirements? How do I integrate that with my specification? [...] Even if it were a dump, in CSV or whatever, you know? Something I could take from here and put there."*
- P7 - *"[...] is there anything implemented today to generate a... [report] Export the data? From the user's point of view?"*
- P9 - *"[...] another thing I was going to ask you was... what is the final document you generate?"*

Two researchers also drew attention to requirements that are not obtained directly from stakeholders, but are inherent to the critical system:

- P3 - *"The problem is that, in modern systems, the stakeholder does not exist. That is why we have been working hard with the idea of conjectural requirements, that is, when you start an innovation, startup, or ideation process, you often do not have a stakeholder, you have an expectation of technology, right? And that is something,*

and you are going to build a certain initial solution, right? And you're going to try to implement that solution, but you have several alternative solutions. You don't have a stakeholder to tell you I want this one or I want that one. So, what do organizations usually do? You've seen these AB tests, right? That is, I transfer the elicitation of requirements, I don't have a stakeholder, I have potential users to whom I will randomly deliver construction options and based on some measure (it's not worth going into detail here), but based on some measure that I will collect, I know whether solution A or solution B is more pleasing to individuals, right? And then I pass it on and take it as my requirement. [...] So, that's another thing you have to look at, because here's the thing: if you close with the expectation that you have to have a stakeholder with you, you may not have a project to develop anymore."

- P4 - *"The customer's requirement is what the customer wants to see. It may not be necessary for you to develop the complete system. You may have requirements inherent to development. [...] So, development requirements for critical systems, the customer's vision alone is not enough. Because on its own, it gives a view of what the system delivers externally, not what it is internally. Internally is where you may have safety and security issues."*

Another point raised in the *SafeSecIoT Canvas* was the name of the Stakeholders building block, since they may not exist in some projects:

- P3 - *"Is it worth keeping this square here with the name stakeholder or putting it as persona, because it is more general? Persona/stakeholder. It's a suggestion.''*

In addition, there was a suggestion to create templates (structured entries) to facilitate the specification of hazards and threats:

- P1 - *"Could you create a template for people to specify hazards, and for people to specify threats, you know¿'*

Still in the context of ensuring quality specifications, it was emphasized that requirements must be precise (unambiguous) and verifiable, which can be supported by the use of standards such as ISO 29148 for standardizing the structure of requirements:

- P4 - *"So, there has to be deterministic precision so that when you read the requirement, the person has no room for interpretation. In fact, this is one of the fundamentals of requirements engineering. The requirement cannot be ambiguous, meaning that if I read it, I find one thing, and if you read it, you find another. So, this is something that applies to all areas of critical systems that require security requirements, which use STPA. All standards talk about having a requirement that is unambiguous and verifiable."*

Another important suggestion was to include the identification of a requirement as functional or non-functional:

- P3 - *"Does this come out as a functional or non-functional requirement? [...] When I export, some of these requirements here will indicate functionality, and other requirements will [indicate] restrictions for building the functionalities."*

Another important point identified for quality improvement was an artifact to support requirements validation:

- P4 - *"So, perhaps, you could look into having another support tool, after you have listed all the requirements, to qualitatively evaluate those requirements. I don't know, a checklist, something like that, right? To give you an idea, all standards, all safety engineering standards, all of them, are requirement-oriented."*

Usability and User Experience

Suggestions were also obtained regarding usability and user experience, mainly in relation to the tool. Among these, participants suggested highlighting the instructions for filling out the form and also making some adjustments:

- P6 - *"Acho que tu poderia, assim, só a questão mesmo de visualização, né? Colocar entre colchetes [as instruções de preenchimento] só para mostrar, entendeu? Pelo menos [pra mim], assim, parece que está preenchido já, entendeu?*
- P4 - *"Então, esse é um primeiro ponto que talvez o seu template precisa de ajuste. Talvez seriam necessidades de desenvolvimento, e não no cliente, especificamente, tá?"*

The improvement suggestion provided by P4 complements the previous one in relation to the consideration of critical system requirements inherent to the system in addition to the requirements obtained from stakeholders.

Another participant highlighted an important aspect of collaborative tools, which is the real-time visualization of what other users are doing in the tool:

- P11 - [...] if you and I are editing this project at the same time, will I be able to see what you are editing here?

In this regard, the tool allows each user's entries and changes to be viewed after they have been made and saved in the database. Future work will focus on improving the tool's functions and support for collaboration.

Integration and Support by Tool

As the evaluation was divided into two parts, the first including conceptual artifacts and the second including the tool, during the first part several specialists mentioned the possibility of creating a tool for the approach. After the tool was presented and the opportunity to use it, some experts made suggestions for specific improvements to some of the tasks supported by it.

One of the suggestions made is the automatic generation of information for STPA-based analysis from the *SafeSecIoT Canvas*:

- P06 - *[...] because I got the idea that when you generated things on the canvas, when you clicked on STPA up there, right, it would generate this list of threats, you know?*

The project planning information consolidated in *SafeSecIoT Canvas* is actually used as input or as a basis for the next stages of analysis. However, in the current version of the tool, this process still depends heavily on the analyst, as there is no complete automation of the analysis process, via an AI assistant, for example, which is being considered as a future task. For now, only system-level constraints and safety and security requirements have AI support for generation from hazards/threats and UCAs, respectively.

Along the same lines, another suggested improvement was the separate generation of requirements from each of the UCAs:

- P10 - *"[...] when you use this option to generate there, right, with AI... you generate from each UCA, right? [...] Yeah, my question is because, from a practical point of view, you would need to review each UCA, right? Because you may not be... agreeing or have some problem with the generation, right? So maybe individual selection would be interesting too.''*

Finally, integration with external tools was suggested for generating new artifacts to support the rest of the system development process. For example, integration with EMF Eclipse was suggested:

- P1 - *"Have you ever thought about doing something with Eclipse? [...] Because then you could integrate it with Papirus UML, for example, you know? You could make an extension of Papirus UML. Papirus is a UML modeling tool that is available in Eclipse. [...] And then you would integrate it with Eclipse, right? You would do it in table format... You could do that or a diagram, right? Depending on what you think is most appropriate. That's for the future, right? So, for your doctorate, what you have achieved so far is more than excellent."*

Control Structure

Although highlighted as a strength of the tool, the step of defining the control structure received some suggestions for improvement, mainly related to its detail in the context of STPA-based analysis:

- P11 - *"If you had a tab where you could at least see the control actions, then it would make sense to put their codes there."*
- P11 - *"I think you should be able to register what a controlled system is and what a controller is. The controller receives external communication."*
- P11 - *"If you can make it navigable there or even show it here, it would be very interesting, at least to show the control actions and feedback."*

The suggestions made by P11 contribute to improvements in the definition stage of the control structure of the *SafeSecRETS* tool. The navigable control structure, for example, was considered but has not yet been implemented due to the complexity and time required to implement this functionality.

P10 also raised a question about the refinement of the control structure:

- P10 - *"What I was thinking, if you need to refine the control structure, add another component and such... [is it possible to do that?]"*

The insertion of new components at the same level of abstraction is supported by the tool. However, refinement, in the sense of creating a new, more detailed version, could be achieved by a new control structure refinement feature, maintaining the more abstract structure developed previously and detailing the components and subsystems from it.

Methodological Considerations

Regarding methodology, two participating experts suggested conducting experimental studies for a more detailed evaluation of the proposal:

- P1 - *"[...] did you conduct usability testing, for example, to collect data on the time it took people to complete the task, the number of errors they made, and perhaps, at the end, give them a form to evaluate usability? There is a questionnaire called the System Usability Scale [...]."*
- P4 - *"[...] we have to conduct experimental evaluations to identify opportunities for improvement. In other words, the approach you will deliver is not the one you defined. The approach you have to deliver is the one you will identify as needing to be developed based on the evaluation."*

- P6 - *"Yes, another thing you can do is apply it to a group of undergraduate students, right? You and your advisor teach an engineering course, and you don't say almost anything, you don't say almost anything, and see how they behave, right? By creating a control group, in this case, you could even take this work you're doing as an experiment and check it, right? Which one had a better specification, you know? Using both techniques. But anyway, it's a hypothesis for the next work, if you want to run [an experiment]."*

Regarding usability and experimental evaluation, it was partially carried out in the context of the controlled experiment presented in the previous chapter, but using UMUX instead of the System Usability Scale (SUS), in addition to NASA TLX and TAM (for the experimental group). UMUX is an alternative to SUS that has been widely adopted in usability evaluation research. However, this evaluation was performed only with *SafeSecIoT Canvas*, and the artifact was used in printed form, before the tool was created. It has not yet been possible to perform a complete usability evaluation with the *SafeSecRETS* tool and the process, which will be included as future work.

The experimental evaluation suggested by P4 and P6 in this same context has not yet been carried out due to time constraints. It is planned as an important future task to refine the proposal presented, and in this context, other more specific aspects such as usability will be evaluated.

In addition to the suggestions for new evaluations, other considerations suggest specific improvements in the methodology for presenting the proposal, such as in the presentation of what was extended in relation to STPA:

- P3 - *"[...] you will have to show [an artifact] that presents the STPA, which shows where the STPA variations were, right? How are you proceeding and what is the result.''*

Another improvement was suggested regarding the creation of some type of diagram to show an overview (summary) of the approach:

- P11 - *"So, a flowchart or an organizational chart or something similar will help you with this step [of showing an overview]."*

9.5 Threats to Validity

In this section, we present an analysis of the threats to the validity of the study. There are different classification schemes for different types of threats to the validity of an experiment [Wohlin, 2014]. The four types of threats considered to address the

reliability of this study and demonstrate the extent to which its results are valid and were not influenced by biases arising from the researchers' perspective were: i) conclusion validity, ii) internal validity, iii) construct validity, and iv) external validity.

9.5.1 Construct Validity

Construct validity concerns the alignment between the theoretical constructs and their operationalization in measurement. In this study, it refers to verifying whether the questionnaires accurately captured PU, PEOU, and ITU, and TTF in the case of the tool.

The descriptive analysis results indicate that the constructs were appropriately defined and applied, consistent with the literature and minimizing divergent interpretations among participants. The main threats to construct validity observed in this study are discussed below:

- *Single method (survey as the only technique)*: relying mainly on questionnaires may limit construct evaluation, as deeper insights could emerge from case studies or real-world experimentation. To mitigate this, we combined quantitative and qualitative analyses, used both Likert-scale and open questions, and incorporated expert feedback during the process.
- *Interaction between constructs and context*: interpretations of "usefulness" or "ease of use" may vary depending on participants' prior experience with RE, STPA, or similar tools, potentially influencing responses beyond the intrinsic quality of the artifacts.
- *Expectation effects (participants or researcher)*: participants may answer in line with perceived expectations, and researchers may risk biased interpretation. This was mitigated by ensuring anonymity and voluntary participation, basing evaluations strictly on collected data, and verifying qualitative coding through independent review by two doctoral students.

9.5.2 Internal Validity

In this study, internal validity refers to the confidence that the evaluation results accurately reflect the impact of the proposed artifacts on participants' perceptions of safety and security RE in critical IoT systems. A potential threat arises from uncontrolled factors (such as prior familiarity with RE methods, expertise in safety/security, or individual preferences) that may have influenced responses. Although the study design aimed to minimize such effects, external variables could still have shaped participants' perceptions of the artifacts.

Next, we analyze the threats to internal validity for the study conducted:

- *History*: possibility of different treatments being applied at different times. To mitigate this threat, all assessments were conducted within a short period of 20 days, scheduled according to participant availability, carried out synchronously and individually, and following a standardized script.
- *Testing*: conducting multiple evaluations may influence subsequent responses. In our study, participants first evaluated the *SafeSecIoT Canvas* model and the *STPA-SafeSecIoT* method, followed by the *SafeSecRETS* tool. To reduce this effect, participants were not informed in advance which artifacts they would evaluate, avoiding expectations about the existence of the tool.
- *Maturation*: changes in participants' attitudes or fatigue over time. To minimize this threat, concise videos and questionnaires were used to standardize and shorten the process. The researcher monitored all sessions to clarify questions, which lasted 70–120 minutes depending on participant pace.
- *Instrumentation*: limitations from the instruments used. This threat was mitigated through prior review of the questionnaires by two doctoral students and the research advisor, who provided improvements and feedback.
- *Selection*: differences in participant characteristics. We selected individuals with strong knowledge in RE and at least one additional area related to the proposal (e.g., safety, security, IoT, or STPA), ensuring diverse perspectives for analysis (see Table 9.1).
- *Experimental mortality*: participant dropout. Two participants were lost due to technical issues with the first questionnaires, but nine valid responses remained, which is considerable for the target audience.
- *Experimenter bias*: potential influence of the researcher. To mitigate this threat, the evaluation was systematized and standardized through recorded videos, scripted sessions, and adherence to reference literature, ensuring clarity, impartiality, and consistency.

9.5.3 External Validity

External validity refers to the extent to which the results of a study can be generalized beyond the specific conditions in which it was conducted. In this research, which involved requirements engineering experts with experience in safety and security of critical IoT systems, the key concern is whether the relationships observed between the proposed artifacts (*SafeSecIoT Canvas*, *STPA-SafeSecIoT*, and *SafeSecRETS*) and participants' perceptions (PU, PEOU, ITU) can be extended to broader contexts. Below we outline the main threats to external validity and the strategies adopted to mitigate them:

- *Sample representativeness*: occurs when participants do not adequately reflect the

target population. Although the sample size does not allow statistical generalization, all participants are RE specialists with expertise in one or more relevant areas. Based on their profiles and the results obtained, we consider them adequate representatives of the intended population.

- *Interaction between selection and treatment*: refers to the risk of including non-representative subjects. To address this, we invited participants with expertise in RE and at least one additional relevant domain (safety, security, IoT, or STPA). Their professional experience was validated through data collection, and invitations were directed to active researchers with extensive background in these fields.
- *Interaction between environment and treatment*: concerns the risk that the study setting and materials do not reflect real-world practice. To mitigate this, we adopted a widely recognized case from the literature—an automatic insulin delivery system. Although simplified for feasibility, the example was based on specialized studies and reviewed by a domain expert. This concise case facilitated understanding of the approach and artifacts while maintaining realism.

9.5.4 Conclusion Validity

Conclusion validity concerns the extent to which inferences about the relationship between treatment and outcomes are statistically sound. In this study, it refers to verifying whether the analyses reliably support a relationship between the use of the proposed artifacts (*SafeSecIoT Canvas*, *STPA-SafeSecIoT*, and *SafeSecRETS*) and participants' perceptions of PU, PEOU, and ITU.

The main threats to validity in the present study include:

- *Low statistical power*: the limited number of participating experts may reduce the ability to detect statistically significant relationships, weakening confidence in the results. To mitigate this, we invited participants with extensive experience in knowledge areas relevant to the proposal.
- *Measurement error*: subjective interpretation of questionnaire items may introduce inconsistencies unrelated to the effect investigated. To address this, assessments were conducted synchronously, allowing participants to ask clarifying questions. Moreover, widely recognized and validated instruments from the literature were used to minimize interpretation issues.
- *Response bias*: participants may provide biased responses for reasons beyond the researcher's control, affecting inference accuracy. To reduce this risk, assessments were carried out by experienced researchers using a standardized script.
- *Data noise*: external or uncontrolled factors (e.g., organizational context or prior familiarity with RE methods) may influence responses. Nonetheless, the descriptive

analysis indicated high questionnaire reliability and low standard deviation, showing consistent perceptions among participants across constructs and artifacts.

9.6 Chapter Summary

This chapter presented the evaluation of the safety and security RE process for critical IoT systems based on the proposed artifacts: i) the *SafeSecIoT Canvas* model and ii) the *STPA-SafeSecIoT* method; and also iii) the *SafeSecRETS* tool, which provides technological support to the process. The process was evaluated using questionnaires based on the Technology Acceptance Model (TAM) and Task-Technology Fit (TTF), with the participation of Requirements Engineering experts with experience in safety/security and/or IoT and/or STPA. The objective of the evaluation was to understand the participants' perceptions of PU, PEOU, ITU, and the overall applicability of the proposal.

The results show a generally positive perception of the conceptual artifacts, further reinforced by the supporting tool. PU scores indicate that experts recognized the potential benefits of applying the artifacts in the RE process for critical IoT systems. Regarding PEOU, the *SafeSecIoT Canvas* model and *SafeSecRETS* tool scored slightly higher than PU, while the *STPA-SafeSecIoT* method scored slightly lower, reflecting a modestly steeper learning curve for the STPA-based method. Consistent with these findings, ITU scores were high for the conceptual artifacts (6.42/7) and even higher for the tool (6.70/7). Overall, these results suggest strong acceptance of the approach among RE specialists, regardless of their experience in safety/security, STPA-based analysis, or IoT systems.

The responses to the open questions and the contributions obtained from the transcripts of the feedback received during the evaluation meetings were analyzed qualitatively and grouped into three categories: Strengths, Limitations, and Suggestions for improvement.

Based on the evaluation stages conducted, it was possible to identify the strengths of the work carried out and also different possibilities for improvement, as well as relevant directions for the continuity of this investigation, some of which are already being incorporated into the continuation of the research. In the following chapter, which concludes this thesis, we present the conclusions of this work, in which we revisit the proposed research questions, highlight the main contributions achieved, discuss the limitations encountered, and indicate perspectives for the continuation of this research and future work.

CHAPTER 10

Conclusions

This chapter presents the final considerations of this thesis, revisiting the research problem and the conclusions obtained from the work carried out. We also present a brief description of the evolution of the research, which discusses the research paths and decisions, as well as the contributions, limitations, and lessons learned at the end of this doctoral work.

10.1 Thesis Overview

The research problem addressed in this thesis concerns the lack of a systematic process and specific artifacts to support RE for safety and security from the early stages of project design. To address this gap, we propose artifacts that support the planning, elicitation, analysis, and specification of requirements for critical IoT systems. These artifacts were implemented in a software tool to enhance both the utility and usability of the proposed approach. A key differentiator of this research is the adoption of a canvas model, which provides structured support for integrated STPA-based safety and security analysis, contributing to a clearer definition of system scope and to the systematization of risk and requirements analysis activities.

After conducting our evaluations, we concluded that the proposed artifacts, the *SafeSecIoT Canvas* model, the *STPA-SafeSecIoT* method, and the *SafeSecRETS* tool, offer a significant contribution to the RE process for critical IoT safety and security systems by integrating their strategic and analytical perspectives. The *SafeSecIoT Canvas* was recognized as a solid starting point, capable of organizing planning and elicitation activities, supporting scope definition, and facilitating communication between stakeholders and technical staff, in addition to reducing barriers to entry for professionals less experienced with the RE process. The *STPA-SafeSecIoT*, on the other hand, was valued for detailing and structuring the analysis, enabling the specification of safety and security requirements in a systematic and integrated manner, an aspect seen as essential in complex cyber-physical systems.

The *SafeSecRETS* tool, in turn, represents an important step toward enabling and implementing the integrated application of the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts in the RE process for critical IoT systems. By offering automated support and facilitating the tasks required in the process, such as filling in information, organizing, and tracking between analysis steps, the tool helps reduce the perceived complexity of manually applying the artifacts, increasing the usefulness and ease of use of the approach and, potentially, the consistency of the analysis. In addition, because it is web-based and collaborative, it allows open and immediate access by all users involved in a project, ensuring that improvements and new versions are made available on an ongoing basis. Thus, *SafeSecRETS* not only materializes the proposed approach in the form of a software tool, but also expands its practical utility by facilitating the adoption of artifacts by different user profiles in real contexts of critical system development.

10.2 Research Evolution

This doctoral research began with the goal of aligning safety and security requirements through an analysis approach based on the STPA method. We proposed an extension of STPA to comprehensively address the safety and security aspects of critical IoT systems, thus providing a systematic basis for requirements specification. However, during the development of the process, it was realized that one of the weaknesses of the direct application of STPA was related to the project preparation stage, in particular the clear definition of the scope of the system to be analyzed and the identification of strategic elements that impact safety and security from the initial stages.

Given this finding, the research evolved to include a planning stage supported by a canvas-format artifact, designed to collect strategic project information, elicit initial requirements, and clarify characteristics specific to critical IoT systems, such as components and actions performed, data and connectivity, relevant assets, and unacceptable potential losses. This information became the basis for guiding the STPA analysis, reducing uncertainties and providing greater clarity about the system context. Thus, the canvas artifact not only complemented the approach but became fundamental to supporting the entire process of analysis and specification of requirements, establishing a link between strategic planning and the detailed application of approaches based on the STPA method.

10.3 Summary of Contributions

The main contributions of this research are summarized below:

- The construction of the *MM4Canvas* metamodel: designed as a methodological support for the development of artifacts in canvas format. This metamodel provides a conceptual and formal structure that allows the systematization of the canvas creation process, ensuring consistency, reuse, and clarity in the definition of its constituent elements [Veiga et al., 2024b].
- The development of the *SafeSecIoT Canvas* model: instantiated from the metamodel, it is an artifact aimed at project planning and requirements elicitation in critical IoT systems. The model was designed to capture strategic project information, specific aspects of IoT systems, as well as initial requirements related to safety and security. In this sense, the artifact assists both in defining the scope and in reducing uncertainties in the process, serving as a structured starting point for more in-depth analysis [Veiga et al., 2024a].
- The development of the *STPA-SafeSecIoT* method: an extension of STPA that enables the joint analysis and specification of safety and security requirements. This adaptation overcame the limitations of the original versions of the method, offering an integrated approach to identify, analyze, and track critical requirements in IoT systems [Veiga and Bulcão Neto, 2023].
- Evaluation with students: with the objective of validating the *SafeSecIoT Canvas* model and analyzing its applicability, even by users with less experience. This study made it possible to verify the clarity, usefulness, and even the pedagogical potential of the artifact.
- The development of the *SafeSecRETS* tool: as a concrete way to integrate the proposed artifacts, the *SafeSecRETS* tool was designed and developed, which implements *SafeSecIoT Canvas* and *STPA-SafeSecIoT* in a single solution. This tool expands the practical applicability of the approach, providing automated support for the RE process, organizing the workflow and information collected at each stage, facilitating traceability, and contributing to reducing the complexity of a manual process [Veiga et al., 2025b].
- Evaluation with experts: with the objective of validating the application of the proposed artifacts and the approach as a whole, initially considering the model and method and, subsequently, the use of the tool. The evaluation provided evidence on the usefulness, ease of use, intended use, and suitability of the artifacts in the context of critical IoT systems. This step was essential to obtain qualified feedback and identify strengths, limitations, and suggestions for improvement for each of the artifacts and for the approach in general.

In summary, the main contribution of the research is the proposal of an integrated approach to support the RE of critical IoT systems from the perspectives of safety and security. By articulating strategic planning, elicitation, analysis, and specification,

and providing a support tool, the research contributes both methodologically and practically, offering means to reduce uncertainties, improve requirements documentation, and facilitate the adoption of more complex techniques such as STPA.

10.4 Limitations

The evaluations carried out revealed limitations that will be addressed in future work, such as the need for greater clarity in the integration flow between artifacts, dependence on the analyst's prior knowledge, and the absence of automated mechanisms for validating requirements. Suggestions for improvement converge on the creation of practical examples, clearer guides, support templates, and technological resources that make the application more fluid. Thus, it is concluded that, although the artifacts demonstrate potential to support the practice of requirements engineering in critical IoT systems, their evolution depends on the incorporation of resources that increase technical rigor and automation, without losing sight of the strategic function of aligning initial planning with the detailed specification of safety and security requirements.

10.5 Lessons Learned

The development of this research made it possible to draw important lessons, which not only reinforced the theoretical and methodological foundations of the work but also provided practical evidence to support its applicability and future evolution:

- *MM4Canvas* metamodel: the construction of a metamodel proved essential to ensure consistency and systematization in the creation of canvas. The prior definition of a conceptual framework reduces ambiguities and facilitates the replicability of the artifact by other researchers and professionals.
- *SafeSecIoT Canvas* model: when designing and applying this artifact, it was realized that the strategic planning phase is decisive for the quality of the subsequent analysis. The involvement of business, design, and IoT context aspects in the initial stage strengthens the elicitation of requirements and contributes to reducing gaps in the following phases.
- *STPA-SafeSecIoT* method: The extension of STPA highlighted the challenge of considering safety and security in a unified analysis. Although methodologically more complex, the effort of this integration results in significant gains in completeness and traceability, allowing this process to evolve into a more in-depth analysis that considers dependencies and potential conflicts between requirements.

- *SafeSecRETS* tool: practical implementation has shown that tool support is decisive for the adoption of an RE approach in real contexts, reducing the perception of complexity and increasing the understanding of its usefulness. Automating processes, organizing information, and providing a software tool in an accessible environment considerably increases the acceptance and usefulness of the approach.
- Evaluation with students: application in the context of an undergraduate course, a learning environment, showed that the approach has potential for teaching safety and security concepts in a structured way and that canvas-based planning is easy to understand and can be easily applied even to a more complex domain. This type of artifact can reduce barriers to entry for more complex techniques, such as STPA, by making the process more visual and intuitive.
- Evaluation with experts: interaction with experienced professionals highlighted the importance of qualitative evaluations to validate the applicability of the proposed artifacts and identify limitations and possible improvements to the approach. Dialogue with experts not only legitimizes the artifacts, highlighting their strengths, but also points to concrete paths for improvement and future evolution.

Overall, the main lesson learned from this work is that the integration of strategic planning, systematic analysis, and tool support strengthens requirements engineering in critical IoT systems, making it more structured, understandable, and applicable in different contexts.

10.6 Future Work

Based on the evaluation conducted with experts, opportunities for further research were identified to strengthen both the approach and the tool developed. Below, we discuss possible future work, classifying it into categories according to its nature.

The first category relates to the quality and standardization of requirements. Suggestions made by experts, such as the classification between functional and non-functional requirements, the prioritization of requirements, the definition of security profiles according to system needs, and the use of standards such as ISO 29148 and ISO 21434, point to the importance of aligning the process with established standards for requirements and security engineering. In addition, the potential for creating validation checklists and templates to guide the steps of the method and the specification of hazards and threats, which could improve the quality of the specifications, was highlighted.

The second category refers to integration and tool support. The need to improve traceability from some points in the analysis and to create new types of relationships, for example, stakeholders with losses, demonstrates the importance of mechanisms for

linking elements of the analysis. Improvements were also suggested in the control structure, detailed visualization of control actions, and support for real-time collaborative editing, reinforcing the central role of the *SafeSecRETS* tool. In addition, the possibility of using frameworks such as EMF Eclipse to facilitate future integrations was pointed out. demonstrates the importance of more robust mechanisms for linking elements of the analysis

Finally, with regard to safety and security, experts highlighted the relevance of expanding the treatment of security through the use of consolidated techniques such as STRIDE and *threat analysis*, as well as highlighting the different security perspectives (confidentiality, integrity, and availability) throughout the analysis process. This evolution would contribute to consolidating the integration between safety and security, expanding the scope of the approach to different domains of critical IoT systems. In addition, based on feedback received from some experts on the scope of the proposed approach, a second future work in this line is to experiment with the approach for different domains of critical systems, even if they are not specifically IoT.

In general, future work points to the consolidation of the proposal in three dimensions: i) strengthening the standardization and quality of requirements; ii) expanding tool support; and iii) covering safety and security aspects. These advances may increase the applicability of the approach in industrial contexts and also stimulate its large-scale adoption in the engineering process of critical IoT system requirements.

Bibliography

[Aguilar-Calderón et al., 2022] Aguilar-Calderón, J.-A., Tripp-Barba, C., Zaldívar-Colado, A., and Aguilar-Calderón, P.-A. (2022). Requirements Engineering for Internet of Things (IoT) Software Systems Development: A Systematic Mapping Study. *Applied Sciences*, 12(15).

[Ajayi et al., 2025] Ajayi, O. O., Alozie, C. E., and Abieba, O. A. (2025). Enhancing cybersecurity in energy infrastructure: strategies for safeguarding critical systems in the digital age. *Trends in Renewable Energy*, 11(2):201–212.

[Ashton et al., 2009] Ashton, K. et al. (2009). That 'internet of things' thing. *RFID journal*, 22(7):97–114.

[Balfe et al., 2017] Balfe, N., Leva, M., Ciarapica-Alunni, C., and O'Mahoney, S. (2017). Total project planning: Integration of task analysis, safety analysis and optimisation techniques. *Safety Science*, 100:216–224.

[Bergenstal et al., 2016] Bergenstal, R. M., Garg, S., Weinzimer, S. A., Buckingham, B. A., Bode, B. W., Tamborlane, W. V., and Kaufman, F. R. (2016). Safety of a Hybrid Closed-Loop Insulin Delivery System in Patients With Type 1 Diabetes. *JAMA*, 316(13):1407–1408.

[Borgia, 2014] Borgia, E. (2014). The Internet of Things vision: Key features, applications and open issues. *Computer Communications*, 54:1–31.

[Cao and Ramesh, 2008] Cao, L. and Ramesh, B. (2008). Agile Requirements Engineering Practices: An Empirical Study. *IEEE Software*, 25(1):60–67.

[Caroli, 2018] Caroli, P. (2018). *Lean Inception: como alinhar pessoas e construir o produto certo*. São Paulo: Editora Caroli, 1st edition.

[Carreras Guzman et al., 2021] Carreras Guzman, N. H., Zhang, J., Xie, J., and Glomsrud, J. A. (2021). A Comparative Study of STPA-Extension and the UFoI-E Method for Safety and Security Co-analysis. *Reliability Engineering & System Safety*, 211:107633.

[Chung and do Prado Leite, 2009] Chung, L. and do Prado Leite, J. C. S. (2009). On non-functional requirements in software engineering. *Conceptual modeling: Foundations and applications: Essays in honor of john mylopoulos*, pages 363–379.

[Conte et al., 2009] Conte, T., Cabral, R., and Travassos, G. H. (2009). Aplicando grounded theory na análise qualitativa de um estudo de observação em engenharia de software–um relato de experiência. In *V Workshop" Um Olhar Sociotécnico sobre a Engenharia de Software"(WOSES 2009)*, volume 1, pages 26–37. sn.

[Corbin and Strauss, 2014] Corbin, J. and Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications.

[Davis, 1989] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3):319–340.

[Deshpande et al., 2019] Deshpande, S., Pinsker, J. E., Zavitsanou, S., Shi, D., Tompot, R., Church, M. M., Andre, C., Doyle, F. J., and Dassau, E. (2019). Design and clinical evaluation of the interoperable artificial pancreas system (iaps) smartphone app: Interoperable components with modular design for progressive artificial pancreas research and development. *Diabetes Technology & Therapeutics*, 21(1):35–43.

[Easterbrook et al., 2008] Easterbrook, S., Singer, J., Storey, M.-A., and Damian, D. (2008). *Selecting Empirical Methods for Software Engineering Research*, pages 285–311. Springer London, London.

[Finocchio-Júnior, 2013] Finocchio-Júnior, J. (2013). *Project Model Canvas*. Elsevier Brasil.

[Finstad, 2010] Finstad, K. (2010). The usability metric for user experience. *Interacting with Computers*, 22(5):323–327. Modelling user experience - An agenda for research and practice.

[Firesmith et al., 2003] Firesmith, D. et al. (2003). Engineering security requirements. *Journal of Object Technology*, 2(1):53–68.

[Friedberg et al., 2017] Friedberg, I., McLaughlin, K., Smith, P., Laverty, D., and Sezer, S. (2017). STPA-SafeSec: Safety and security analysis for cyber-physical systems. *Journal of Information Security and Applications*, 34:183–196.

[Garro et al., 2019] Garro, A., Vaccaro, V., Dutré, S., and Stegen, J. (2019). Cyber-physical systems engineering: model-based solutions. In *Proceedings of the 2019 Summer Simulation Conference*, pages 1–12.

[Glaser and Strauss, 1967] Glaser, B. and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Observations (Chicago, Ill.). Aldine.

[Glass, 1994] Glass, R. (1994). The software-research crisis. *IEEE Software*, 11(6):42–47.

[Glomsrud and Xie, 2019] Glomsrud, J. A. and Xie, J. (2019). A structured STPA safety and security co-analysis framework for autonomous ships. In *European Safety and Reliability conference, Germany, Hannover*.

[Gomola and Bouwer Utne, 2024] Gomola, A. and Bouwer Utne, I. (2024). A novel stpa approach to software safety and security in autonomous maritime systems. *Heliyon*, 10(10).

[Goodhue and Thompson, 1995] Goodhue, D. L. and Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 19(2):213–236.

[Goodrich and Tamassia, 2011] Goodrich, M. T. and Tamassia, R. (2011). *Introduction to computer security*. Pearson London, UK.

[Greer et al., 2019] Greer, C., Burns, M., Wollman, D., and Griffor, E. (2019). Cyber-Physical Systems and Internet of Things.

[Gromule et al., 2017] Gromule, V., (Jackiva), I. Y., and Pēpulis, J. (2017). Safety and Security of Passenger Terminal: the Case Study of Riga International Coach Terminal. *Procedia Engineering*, 178:147–154.

[Gunes et al., 2014] Gunes, V., Peter, S., Givargis, T., and Vahid, F. (2014). A survey on concepts, applications, and challenges in cyber-physical systems. *KSII Transactions on Internet and Information Systems (TIIS)*, 8(12):4242–4268.

[Guth et al., 2016] Guth, J., Breitenbücher, U., Falkenthal, M., Leymann, F., and Reinfurt, L. (2016). Comparison of IoT platform architectures: A field study based on a reference architecture. In *2016 Cloudification of the Internet of Things (CIoT)*, pages 1–6.

[Hart and Staveland, 1988] Hart, S. G. and Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In Hancock, P. A. and Meshkati, N., editors, *Human Mental Workload*, volume 52 of *Advances in Psychology*, pages 139–183. North-Holland.

[Henderson-Sellers, 2011] Henderson-Sellers, B. (2011). Bridging metamodels and ontologies in software engineering. *Journal of Systems and Software*, 84(2):301–313.

[Heravi et al., 2015] Heravi, A., Coffey, V., and Trigunarsyah, B. (2015). Evaluating the level of stakeholder involvement during the project planning processes of building projects. *International Journal of Project Management*, 33(5):985–997.

[Hevner et al., 2004] Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *MIS quarterly*, pages 75–105.

[Hwang, 2008] Hwang, S. (2008). Utilizing qualitative data analysis software: A review of atlas. ti. *Social science computer review*, 26(4):519–527.

[Inayat et al., 2015] Inayat, I. et al. (2015). A systematic literature review on agile requirements engineering practices and challenges. *Computers in Human Behavior*, 51:915–929.

[ISO/IEC/IEEE, 2023] ISO/IEC/IEEE (2023). Iso/iec/ieee international standard - systems and software engineering – system life cycle processes. *ISO/IEC/IEEE 15288:2023(E)*, pages 1–128.

[Jarzębowicz and Weichbroth, 2021] Jarzębowicz, A. and Weichbroth, P. (2021). A Systematic Literature Review on Implementing Non-functional Requirements in Agile Software Development: Issues and Facilitating Practices. In *Lean and Agile Software Development*, pages 91–110. Springer International Publishing.

[Kaleem et al., 2019] Kaleem, S., Ahmad, S., Babar, M., Akre, V., Raian, A., and Ullah, F. (2019). A Review on Requirements Engineering for Internet of Things (IoT) Applications. In *2019 Sixth HCT Information Technology Trends (ITT)*, pages 269–275.

[Karlstrom et al., 2002] Karlstrom, D., Runeson, P., and Wohlin, C. (2002). Aggregating viewpoints for strategic software process improvement—a method and a case study. *IEE Proceedings-Software*, 149(5):143–152.

[Kavallieratos et al., 2020a] Kavallieratos, G., Katsikas, S., and Gkioulos, V. (2020a). Cybersecurity and Safety Co-Engineering of Cyberphysical Systems—A Comprehensive Survey. *Future Internet*, 12(4).

[Kavallieratos et al., 2020b] Kavallieratos, G., Katsikas, S., and Gkioulos, V. (2020b). SafeSec Tropos: Joint security and safety requirements elicitation. *Computer Standards & Interfaces*, 70:103429.

[Kerzner, 2002] Kerzner, H. (2002). *Strategic planning for project management using a project management maturity model*. John Wiley & Sons.

[Khan and Madnick, 2022] Khan, S. and Madnick, S. (2022). Protecting Chiller Systems from Cyberattack Using a Systems Thinking Approach. *Network*, 2(4):606–627.

[Kitchenham and Charters, 2007] Kitchenham, B. and Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering.

[Knight, 2002] Knight, J. C. (2002). Safety critical systems: challenges and directions. In *Proceedings of the 24th International Conference on Software Engineering*, ICSE '02, page 547–550, New York, NY, USA. Association for Computing Machinery.

[Kopetz, 2011] Kopetz, H. (2011). *Internet of Things*, pages 307–323. Springer US, Boston, MA.

[Kriaa et al., 2015] Kriaa, S., Pietre-Cambacedes, L., Bouissou, M., and Halgand, Y. (2015). A survey of approaches combining safety and security for industrial control systems. *Reliability Engineering & System Safety*, 139:156–178.

[Kudo et al., 2020] Kudo, T. N., Bulcão Neto, R. F., and Vincenzi, A. M. R. (2020). Requirement patterns: A tertiary study and a research agenda. *IET Software*, 14(1):18–26.

[Lautieri et al., 2005] Lautieri, S., Cooper, D., and Jackson, D. (2005). SafSec: Commonalities Between Safety and Security Assurance. In *Safety-critical Systems Symposium*.

[Leveson, 1995] Leveson, N. G. (1995). *Safeware: System Safety and Computers*. Association for Computing Machinery, New York, NY, USA.

[Leveson, 2016] Leveson, N. G. (2016). *Engineering a safer world: Systems thinking applied to safety*. The MIT Press.

[Leveson and Thomas, 2018] Leveson, N. G. and Thomas, J. (2018). STPA Handbook. *https://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf*.

[Li et al., 2015] Li, S., Xu, L. D., and Zhao, S. (2015). The internet of things: a survey. *Information systems frontiers*, 17:243–259.

[Lim et al., 2018] Lim, T.-Y., Chua, F.-F., and Tajuddin, B. B. (2018). Elicitation Techniques for Internet of Things Applications Requirements: A Systematic Review. In *Proceedings of the 2018 VII International Conference on Network, Communication and Computing*, ICNCC 2018, page 182–188, New York, NY, USA. Association for Computing Machinery.

[Line et al., 2006] Line, M. B., Nordland, O., Røstad, L., and Tøndel, I. A. (2006). Safety vs. Security? (PSAM-0148). In *Proceedings of the Eighth International Conference on Probabilistic Safety Assessment & Management (PSAM)*. ASME Press.

[Lisova et al., 2019a] Lisova, E., Šljivo, I., and Čaušević, A. (2019a). Safety and security co-analyses: A systematic literature review. *IEEE Systems Journal*, 13(3):2189–2200.

[Lisova et al., 2019b] Lisova, E., Šljivo, I., and Čaušević, A. (2019b). Safety and security co-analyses: A systematic literature review. *IEEE Systems Journal*, 13(3):2189–2200.

[López-Fernández et al., 2015] López-Fernández, J. J., Cuadrado, J. S., Guerra, E., and De Lara, J. (2015). Example-driven meta-model development. *Software & Systems Modeling*, 14(4):1323–1347.

[Lyu et al., 2019] Lyu, X., Ding, Y., and Yang, S.-H. (2019). Safety and security risk assessment in cyber-physical systems. *IET Cyber-Physical Systems: Theory & Applications*, 4(3):221–232.

[Mailloux et al., 2019] Mailloux, L. O., Span, M., Mills, R. F., and Young, W. (2019). A Top Down Approach for Eliciting Systems Security Requirements for a Notional Autonomous Space System. In *2019 IEEE International Systems Conference (SysCon)*, pages 1–7.

[Mellado et al., 2010] Mellado, D. et al. (2010). A systematic review of security requirements engineering. *Computer Standards & Interfaces*, 32(4):153–165.

[Menezes et al., 2021] Menezes, R. V., Sampaio, S., and Marinho, M. (2021). Engenharia de Requisitos Ágil: Extensão de uma Revisão Sistemática da Literatura. In *Anais do WER 2021 - Workshop em Engenharia de Requisitos*.

[Minerva et al., 2015] Minerva, R., Biru, A., and Rotondi, D. (2015). Towards a definition of the Internet of Things (IoT). *IEEE Internet Initiative*, 1(1):1–86.

[Miorandi et al., 2012] Miorandi, D., Sicari, S., De Pellegrini, F., and Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. *Ad Hoc Networks*, 10(7):1497–1516.

[Nawrocki et al., 2014] Nawrocki, J. et al. (2014). Agile Requirements Engineering: A Research Perspective. In *SOFSEM: Theory and Practice of Computer Science*, pages 40–51. Springer International Publishing.

[Nguyen-Duc et al., 2019] Nguyen-Duc, A., Khalid, K., Shahid Bajwa, S., and Lønnestad, T. (2019). Minimum Viable Products for Internet of Things Applications: Common Pitfalls and Practices. *Future Internet*, 11(2).

[OMG, 2002] OMG (2002). *Meta Object Facility (MOF) Specification, Version 1.4.* OMG, Inc.

[Osterwalder and Pigneur, 2010] Osterwalder, A. and Pigneur, Y. (2010). *Business Model Generation: a handbook for visionaries, game changers, and challengers*, volume 1. John Wiley & Sons.

[Pargaonkar, 2023] Pargaonkar, S. (2023). Synergizing requirements engineering and quality assurance: A comprehensive exploration in software quality engineering. *International Journal of Science and Research (IJSR)*, 12(8):2003–2007.

[Peffers et al., 2007] Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3):45–77.

[Pfleeger et al., 2015] Pfleeger, C., Pfleeger, S., and Margulies, J. (2015). *Security in Computing*. Pearson Education.

[Pimentel et al., 2020] Pimentel, M., Filippo, D., and dos Santos, T. M. (2020). Design science research: pesquisa científica atrelada ao design de artefatos. *RE@ D-Revista de Educação a Distância e eLearning*, 3(1):37–61.

[Pinheiro, 2004] Pinheiro, F. A. (2004). Requirements traceability. *Perspectives on software requirements*, pages 91–113.

[Rajkumar et al., 2010] Rajkumar, R., Lee, I., Sha, L., and Stankovic, J. (2010). Cyber-physical systems: The next computing revolution. In *Design Automation Conference*, pages 731–736.

[Ribeiro and Castro, 2022] Ribeiro, Q. and Castro, J. (2022). Safety & Security Alignment in Requirements Engineering Process for Autonomous Vehicles. In *Anais do WER 2022 - Workshop em Engenharia de Requisitos*.

[Rodrigues da Silva, 2015] Rodrigues da Silva, A. (2015). Model-driven engineering: A survey supported by the unified conceptual model. *Computer Languages, Systems & Structures*, 43:139–155.

[Rosacker and Rosacker, 2010] Rosacker, K. M. and Rosacker, R. E. (2010). Information technology project management within public sector organizations. *Journal of Enterprise Information Management*, 23(5):587–594.

[Rose et al., 2015] Rose, K., Eldridge, S., and Chapin, L. (2015). The internet of things: An overview. *The internet society (ISOC)*, 80:1–50.

[Ross et al., 2019] Ross, R., Pillitteri, V., Graubart, R., Bodeau, D., and McQuaid, R. (2019). Developing cyber resilient systems: a systems security engineering approach. Technical report, National Institute of Standards and Technology.

[Sadvandi et al., 2012] Sadvandi, S., Chapon, N., and Piètre-Cambacédès, L. (2012). Safety and Security Interdependencies in Complex Systems and SoS: Challenges and

Perspectives. In Hammami, O., Krob, D., and Voirin, J.-L., editors, *Complex Systems Design & Management*, pages 229–241, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Schmidt, 2006] Schmidt, D. (2006). Guest editor's introduction: Model-driven engineering. *Computer*, 39(2):25–31.

[Shostack, 2014] Shostack, A. (2014). *Threat modeling: Designing for security*. John Wiley & Sons.

[Silva et al., 2021] Silva, D. V., Souza, B. P., Gonçalves, T. G., and Travassos, G. H. (2021). A Requirements Engineering Technology for the IoT Software Systems. *JSERD*, 9(1):11:1 – 11:18.

[Silva et al., 2020] Silva, D. V. d., Gonçalves, T. G., and Travassos, G. H. (2020). A technology to support the building of requirements documents for IoT software systems. In *19th Brazilian Symposium on Software Quality*, SBQS'20, New York, NY, USA. Association for Computing Machinery.

[Sommerville, 2016] Sommerville, I. (2016). *Software Engineering*. Pearson, 10th edition.

[Spanoudakis and Zisman, 2005] Spanoudakis, G. and Zisman, A. (2005). *Software Traceability: A Roadmap*, pages 395–428. World Scientific.

[Takeuchi et al., 2022] Takeuchi, H., Ito, Y., and Yamamoto, S. (2022). Method for Constructing Machine Learning Project Canvas Based on Enterprise Architecture Modeling. *Procedia Computer Science*, 207:425–434.

[Thiée, 2021] Thiée, L.-W. (2021). A systematic literature review of machine learning canvases. INFORMATIK 2021.

[Veiga et al., 2025a] Veiga, E., Kudo, T., and Bulcão-Neto, R. (2025a). SafeSecRETS: A Safety and Security Requirements Tool for Critical IoT Systems. In *Anais do XXXIX Simpósio Brasileiro de Engenharia de Software*, pages 865–871, Porto Alegre, RS, Brasil. SBC.

[Veiga, 2023] Veiga, E. F. (2023). Uma Abordagem para Alinhamento de Requisitos de Segurança e Proteção de Sistemas IoT Críticos. In *Anais do XXVI Congresso Ibero-Americano em Engenharia de Software*.

[Veiga and Bulcão Neto, 2023] Veiga, E. F. and Bulcão Neto, R. d. F. (2023). Toward a Method for Safety and Security Requirements Alignment in Critical IoT Systems. In *Proceedings of the XXXVII Brazilian Symposium on Software Engineering*, SBES '23, page 452–457.

[Veiga and Bulcão-Neto, 2022] Veiga, E. F. and Bulcão-Neto, R. F. (2022). Engenharia de Requisitos de Sistemas IoT e Ciber-Físicos: Resultados Preliminares. In *Anais do WER22 - Workshop em Engenharia de Requisitos*.

[Veiga et al., 2024a] Veiga, E. F., Kudo, T. N., and Bulcão Neto, R. F. (2024a). Linking Agile Planning and Safety and Security Analysis in Critical IoT Systems: An Approach based on ISO/IEC/IEEE 15288. In *Proceedings of the XXIII Brazilian Symposium on Software Quality*, SBQS '24, page 81–91. ACM.

[Veiga et al., 2024b] Veiga, E. F., Kudo, T. N., and Bulcão-Neto, R. F. (2024b). A Canvas Metamodel to Bridging Agile Project Planning and Requirements Engineering. In *Proceedings of the Workshop on Requirements Engineering (WER '24)*, page 1–14.

[Veiga et al., 2025b] Veiga, E. F., Lima, K. L. S., Kudo, T. N., and Bulcão-Neto, R. F. (2025b). SafeSecRETS: A Project Planning Tool for Critical IoT Systems. In *Proceedings of the Workshop on Requirements Engineering (WER '25)*, page 1–8.

[Venkatesh et al., 2003] Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3):425–478.

[Wiesner et al., 2015] Wiesner, S., Hauge, J. B., and Thoben, K.-D. (2015). Challenges for Requirements Engineering of Cyber-Physical Systems in Distributed Environments. In *Advances in Production Management Systems: Innovative Production Management Towards Sustainable Growth*, pages 49–58.

[Wiesner et al., 2017] Wiesner, S., Marilungo, E., and Thoben, K.-D. (2017). Cyber-Physical Product-Service Systems – Challenges for Requirements Engineering. *International Journal of Automation Technology*, 11(1):17–28.

[Wohlin, 2014] Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *18th International Conference on Evaluation and Assessment in Software Engineering, EASE '14, London, England, United Kingdom, May 13-14, 2014*, pages 38:1–38:10.

[Wolf and Serpanos, 2018] Wolf, M. and Serpanos, D. (2018). Safety and Security in Cyber-Physical Systems and Internet-of-Things Systems. *Proceedings of the IEEE*, 106(1):9–20.

[Xing, 2021] Xing, L. (2021). Cascading failures in internet of things: Review and perspectives on reliability and resilience. *IEEE Internet of Things Journal*, 8(1):44–64.

[Yang and Qu, 2016] Yang, Z. L. and Qu, Z. (2016). Quantitative maritime security assessment: a 2020 vision. *IMA Journal of Management Mathematics*, 27(4):453–470.

[Young and Leveson, 2013] Young, W. and Leveson, N. (2013). Systems Thinking for Safety and Security. In *Proceedings of the 29th Annual Computer Security Applications Conference*, ACSAC '13, page 1–8, New York, NY, USA. Association for Computing Machinery.

[Young and Leveson, 2014] Young, W. and Leveson, N. G. (2014). An Integrated Approach to Safety and Security Based on Systems Theory. *Commun. ACM*, 57(2):31–35.

[Yu et al., 2021] Yu, J., Wagner, S., and Luo, F. (2021). Data-flow-based adaption of the system-theoretic process analysis for security (STPA-sec). *PeerJ Computer Science*, 7:e362.

[Zhou et al., 2021] Zhou, X.-Y., Liu, Z.-J., Wang, F.-W., and Wu, Z.-L. (2021). A system-theoretic approach to safety and security co-analysis of autonomous ships. *Ocean Engineering*, 222:108569.

[Zwikael et al., 2014] Zwikael, O., Pathak, R., Singh, G., and Ahmed, S. (2014). The moderating effect of risk on the relationship between planning and success. *International Journal of Project Management*, 32(3):435–441.

APPENDIX A

Evaluation with Undergraduate Students: Additional Information

A.1 Experiment Execution and Data Collection

The experiment, in the context of the IoT course, was organized with the expertise and collaboration of the professor responsible for class 2024/2. It was decided, together with this professor, that the experiment would take place during one of the classes, at the end of the academic semester. This way, students would already have been introduced to all the theoretical content planned for the course and would also be at an advanced stage in relation to the final project for the course (which consisted of developing a real IoT system). This planning was carried out in such a way as not to cause any disruption to the course schedule and also to make the most of the students' training in relation to IoT systems: both from the theoretical content and from practical experience in a project.

A.1.1 Experiment Environment

The experiment was conducted in a workshop format during the morning session on November 6, 2024, from 9:30 a.m. to 12:00 p.m., with a total duration of 2 hours and 30 minutes, including all activities. This period related to the experiment was divided into 5 previously planned parts: leveling, presentation of the application domain, division of groups, experimental activity, and completion of evaluation questionnaires. The steps are detailed below.

A.1.2 Detailed Description of the Stages

Stage 1: Leveling of participants (25 minutes)

The 23 participants who volunteered to take part in the experiment underwent an initial leveling stage, where they were introduced to the concepts of project planning

and critical IoT systems. The purpose of the leveling was to give all participants the same contextual information regarding the planning of IoT system projects with safety and security requirements.

Step 2: presentation of the application domain (10 minutes)

After the initial leveling step, the application domain to be worked on in the experiment was presented. The scenario defined was Automatic Insulin Delivery, which is used to treat patients with type 2 diabetes and can be performed through a critical IoT system. This system contains components such as a continuous glucose monitor (with a wearable sensor that takes readings) and an insulin pump (where the infusion mechanism, the actuator of this system, is located). The same general information about the application domain was presented to all participants, informing them that this would be the system to be worked on in the context of the experiment, but without specifying the task to be performed.

Step 3: defining groups and presenting the activity (5 minutes)

Teams of three participants were drawn at random (one team had two members, due to the total number of 23 participants). Next, a draw was held among the teams formed to divide them into two groups (experimental and control). Thus, 12 participants (four teams of three members) were defined as the experimental group, and 11 participants (three teams of three members and one team of two members) were defined as the control group. The two groups were then allocated to different rooms so that the task to be performed for the experiment could be presented to each group separately and carried out without any influence or bias between the experimental and control groups.

The activity presented to both groups was the same, with only the treatment for its execution being changed. The task to be performed by the groups consisted of creating a project plan for the Automatic Insulin Delivery system, in order to identify the objective, scope, and main requirements of this system, as well as its IoT components and safety and security risks. The treatment requested of the experimental group was the use of a canvas artifact for planning critical IoT system projects, the *SafeSecIoT Canvas*: the research product evaluated during this experiment. The control group received no treatment, meaning that the participants/teams were free to choose the artifact and/or type of document to be developed and delivered.

Step 4: conducting the experimental activity (80 minutes)

After dividing the groups into different work environments (each group worked in a separate room on the third floor of the Baru Class Center at UFG) and presenting

the task (to each group separately), the teams had up to 80 minutes to complete the proposed activity. Participants were instructed that teams that completed the proposed activity before the end of this period could proceed to the next stage of the experiment (completing the evaluation questionnaires).

The experimental group, having a canvas defined as the artifact to be delivered, received the template of the respective artifact printed in A0 format (110cm x 85cm) to be filled out. In addition, post-it notes in different colors (to be used to fill in the information in each field of the canvas) and pens to fill out the post-it notes were provided. This format was chosen because it is widely used for work involving this type of artifact. Meanwhile, in the control group (which did not have an artifact defined as treatment), participants opted to use notebooks to create text documents (usually shared on Google Drive), as well as paper and pens, used for notes and drafts.

Observing the work of the groups, it was noticed that the printed canvas format (used by the experimental group), even without any intervention or motivation from the researcher in this regard, encouraged teamwork and collaboration among all participants in the work, as well as discussion of the system being planned. In the control group, there were also interactions between team members, but these were mainly observed at the beginning of the work. After the initial discussions and definition of the artifact and work methodology (all groups opted to use a shared text document), there was a greater propensity for individual work among the participants and less interaction in the discussions about the system by some members, while others filled out the document.

Throughout the experimental activity, the researcher in charge alternated his presence between the two rooms to observe the work performed by both the control group and the experimental group. The researcher's participation was primarily for observation, interacting with participants only when asked to answer any questions that could be answered, assessing beforehand whether there would be any influence or bias for the participants in the experiment.

Step 5: Completing the evaluation questionnaires (10 minutes)

After the period allocated for completing the task (or as soon as the team finished developing the project plan), all participants were asked to individually answer the evaluation/perception questionnaires. As shown in Figure 7.1, both the experimental and control group participants underwent a workload assessment for subsequent comparison and analysis of the results and verification of the hypotheses. In addition, participants in the experimental group answered the questionnaire based on TAM and UMUX to evaluate the *SafeSecIoT Canvas* artifact, and all participants answered additional open-ended questions.

In addition to the NASA TLX questionnaire, participants in the experimental group, who used the *SafeSecIoT Canvas* artifact, answered a second evaluation questionnaire with objective questions based on TAM and UMUX in order to capture participants' perceptions of the artifact's usefulness, ease of use, and usability. To conclude the evaluation, both groups answered complementary open-ended questions about the activity, which will be presented and discussed later in this chapter.

A.2 Workload Calculation

In order to calculate the perceived workload in relation to the task using NASA TLX, participants had to evaluate the respective task in relation to different aspects, systematized through six parameters (or metrics), namely: i) Mental Demand, ii) Physical Demand; iii) Temporal Demand; iv) Performance; v) Effort; and vi) Frustration. Figure A.1 presents the NASA-TLX questionnaire.

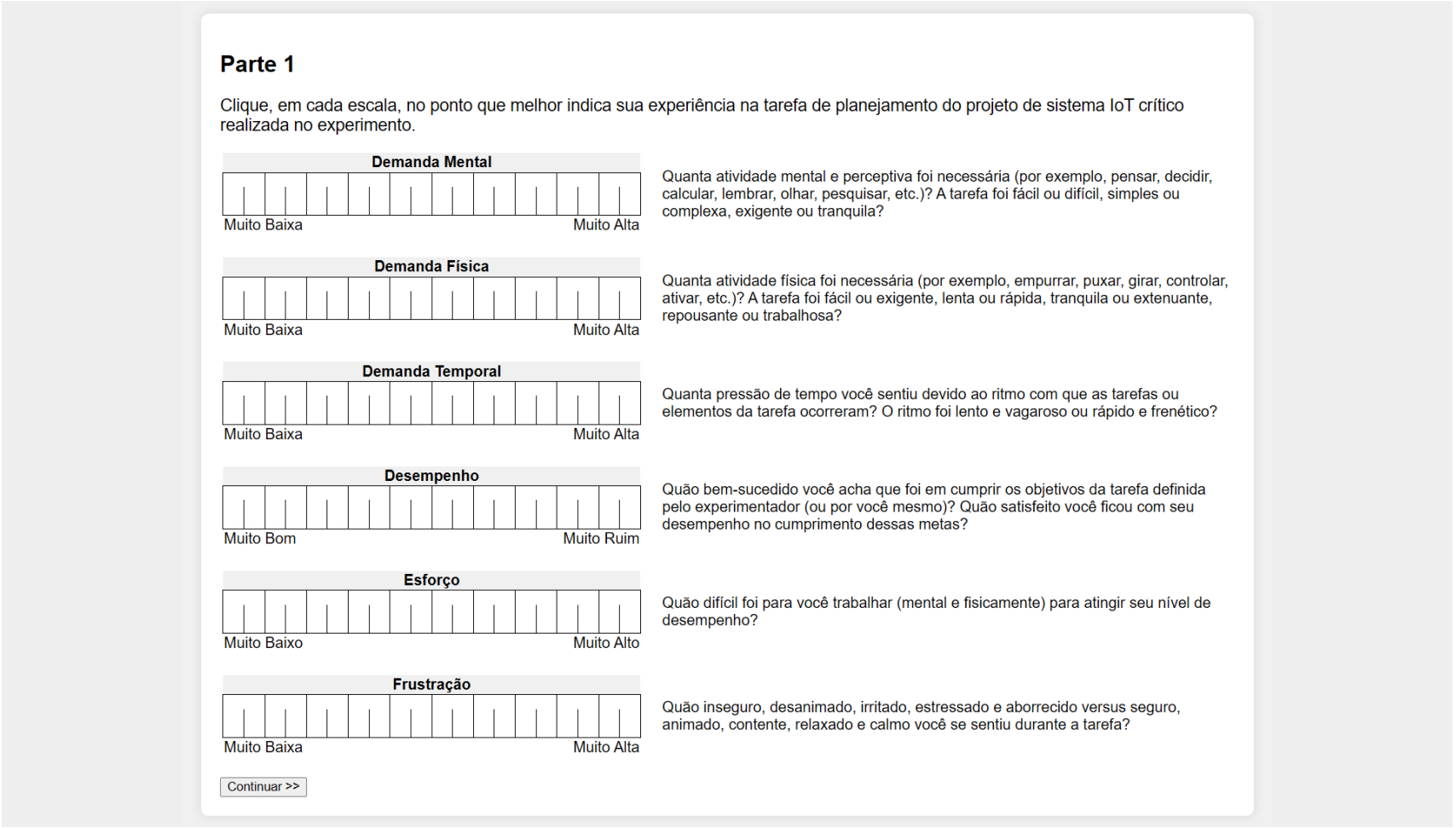


Figure A.1: NASA TLX questionnaire (applied in Portuguese).

Each of the parameters presented was evaluated using a standardized NASA TLX scale, which is divided into 20 parts, representing a score (or percentage) ranging from 5 to 100 for each of the parameters (e.g., 5, 10, 15, 20... 85, 90, 95, 100). An example of this scale is shown in Figure A.2:

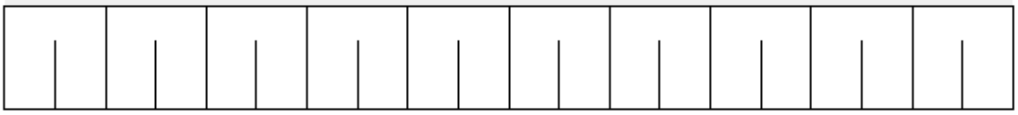


Figure A.2: Standardized NASA TLX scale.

For each aspect evaluated, the left end of the scale (where it starts with the lowest values, e.g., 5, 10, 15...) represents the rating "Very Low," and the right end (where it ends with the highest values, e.g., 85, 90, 100) represents the rating "Very High." The exception is the "Performance" parameter, whose ratings need to be inverted so that the workload is calculated correctly. Thus, the left side represents "Very Good" (with the lowest values) and the right side represents "Very Poor" (with the highest values). Thus, a participant's workload will be a value ranging from 5 to 100.

After the participant assigns the scores related to their perception for each parameter, the second part of the NASA TLX questionnaire consists of assessing the importance that the participant attributes to each parameter for the workload. For this, a direct comparison of each parameter with the others is presented (e.g., Mental Demand or Temporal Demand, Performance or Frustration, etc.). This provides a weight for each parameter, which is used to calculate the weighted average of the scores for all parameters, arriving at the value of the workload perceived by that participant for performing the task.

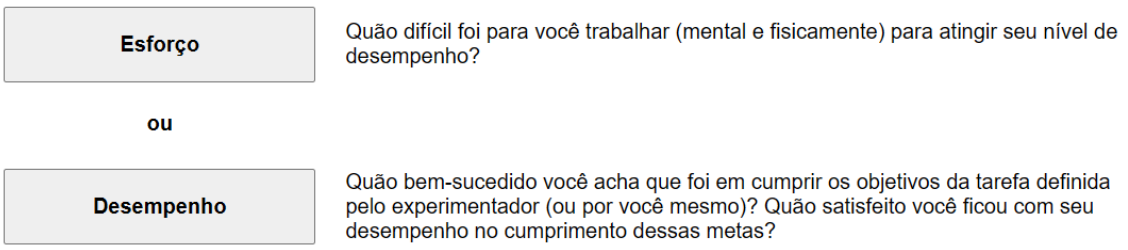


Figure A.3: Example of pairwise comparison between NASA TLX parameters.

Figure A.4 shows the workload result for one of the research participants, to illustrate the data collected[1]:

[1]Note: the classification for the "Performance" parameter is inverse in relation to the others, i.e., the lower the measured value, the better the evaluation.

	Avaliação	Contagem	Peso
Demanda Mental	65	5	0.3333333333333333
Demanda Física	5	1	0.06666666666666667
Demanda Temporal	50	2	0.13333333333333333
Desempenho	15	4	0.26666666666666666
Esforço	40	3	0.2
Frustração	15	0	0

Avaliação Geral da Carga de Trabalho = 40.666666666666664

Figure A.4: Example of workload assessment results for a participant.

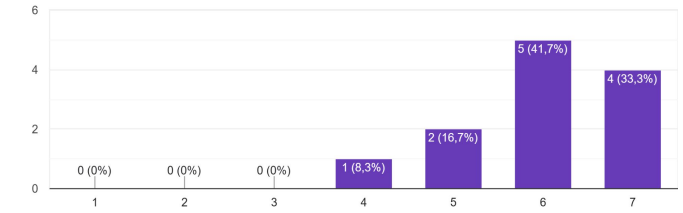
In the example in question, the participant assigned scores to all parameters, which are shown in the "Evaluation" column. In this case, "Mental Demand" was the parameter that received the highest score (i.e., it was the one that demanded the most from the participant) and "Physical Demand" received the lowest score. Next, in the "Count" column, we can see the number of times each parameter was considered most important to the participant in relation to its comparison with the others. In this case, "Mental Demand" was considered most important in all 5 comparisons involving it (each parameter is compared only once with the other 5, totaling 15 comparisons so that all parameters are compared with each other). The count for each parameter is then used to define the "Weight" column, where the sum of the weights must be equal to 1. Next, the weighted average is calculated using the scores for each parameter and the weights obtained: in this case, the workload perceived by the participant was 40.66, on a scale of 5 to 100(%).

A.3 Overview of Results: TAM and UMUX

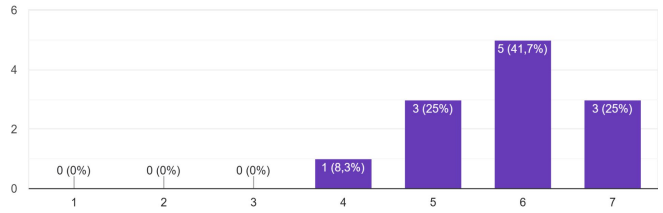
Figures A.5, A.6, and A.7 summarize the responses of the 12 participants in the experimental group regarding the *SafeSecIoT Canvas* artifact. The questionnaire followed the standardized structure of the items that make up the TAM (PU and PEOU) and UMUX constructs. However, they were not presented to participants in sequence or organized by construct. This decision was made in order to adjust the presentation of the questionnaire to students based on the view of what needs to be answered about the artifact. In addition, the third item of the UMUX is identical to the last item of PEOU-TAM ("I consider the *SafeSecIoT Canvas* easy to use") and, therefore, the question was not duplicated for participants.

Figure A.5: Survey Part 1 – Using the *SafeSecIoT Canvas*

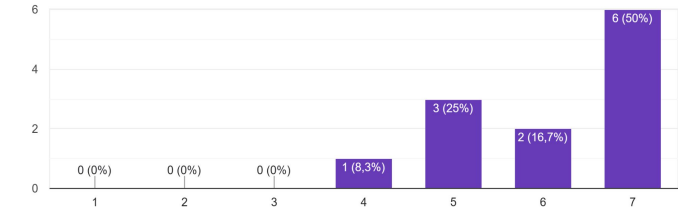
a) me permite realizar tarefas mais rápido.
12 respostas



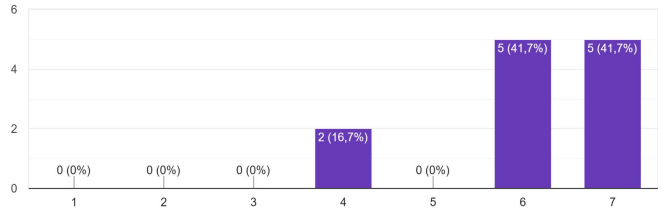
b) melhora meu desempenho no trabalho.
12 respostas



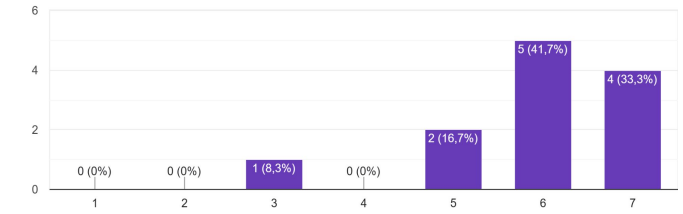
c) aumenta minha produtividade.
12 respostas



d) aumenta minha eficácia no trabalho.
12 respostas



e) torna o meu trabalho mais fácil.
12 respostas



f) é uma experiência frustrante.
12 respostas

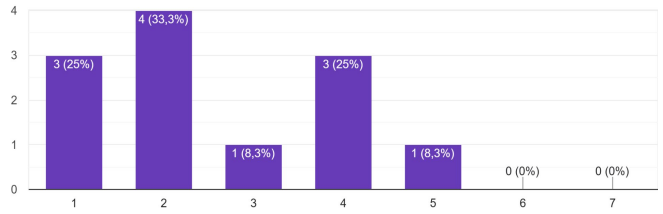


Figure A.6: Survey Part 2 – I consider the *SafeSecIoT Canvas*

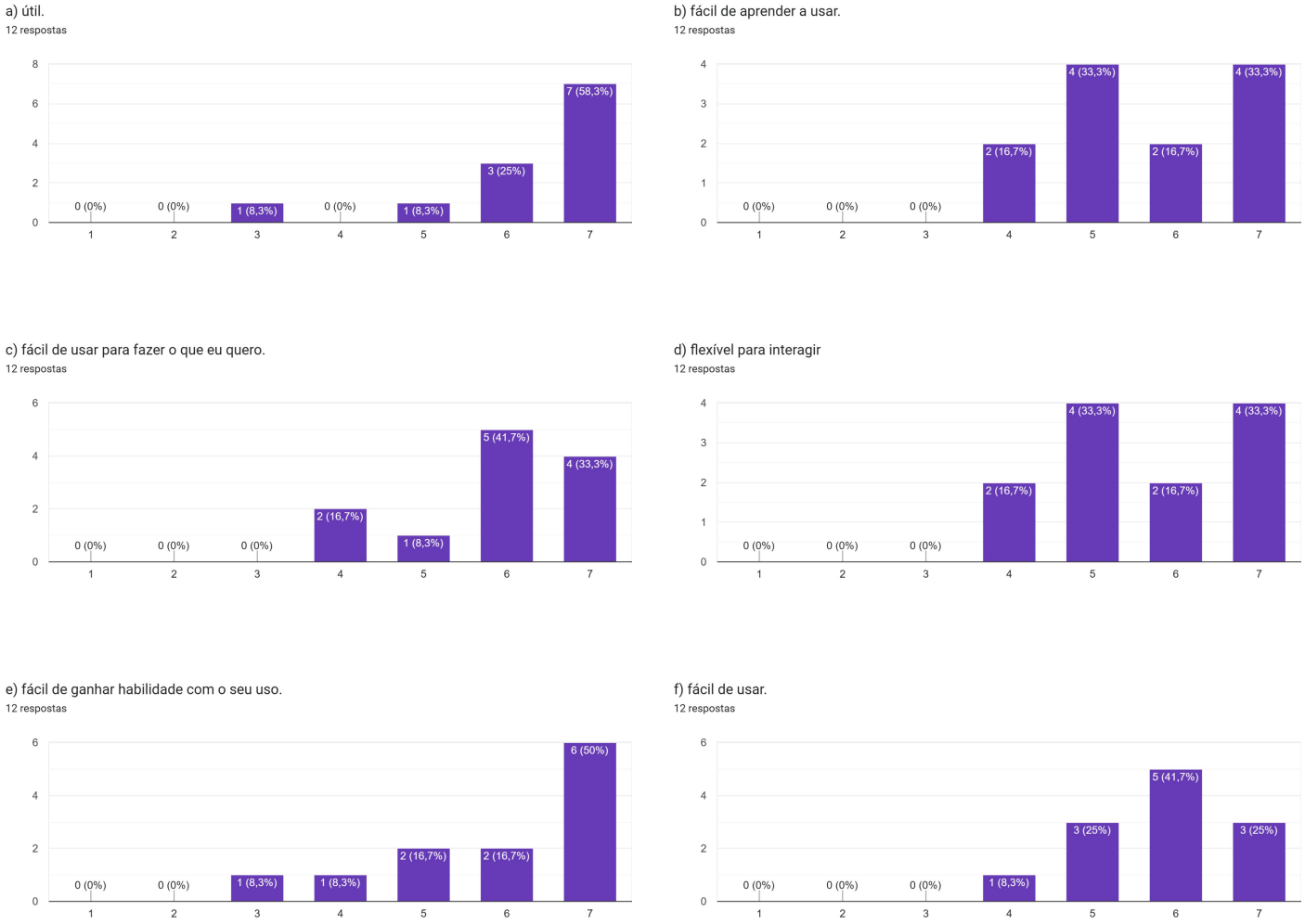
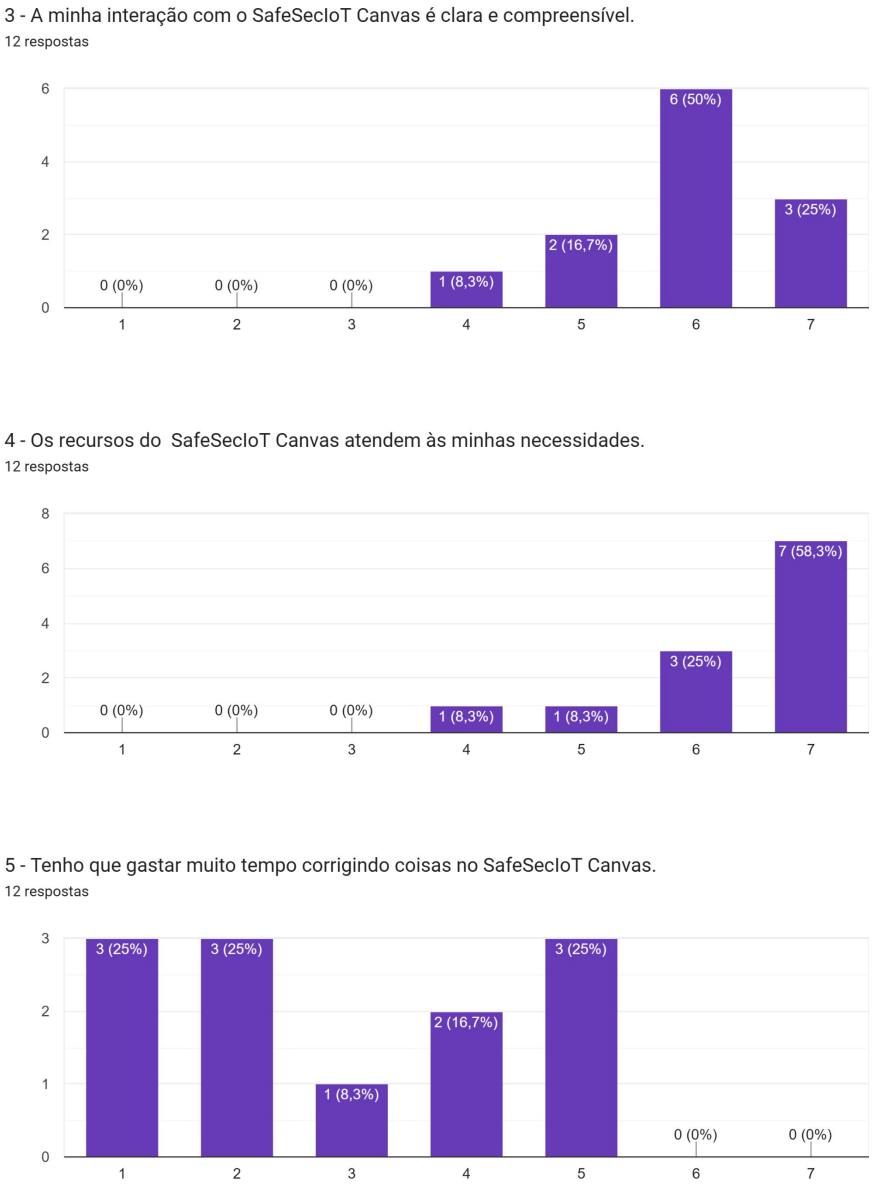


Figure A.7: Survey Part 3 – About the *SafeSecIoT Canvas*



APPENDIX B

Evaluation with Experts: Additional Information

B.1 Overview of TAM and TTF Questionnaires

B.1.1 Constructs and items

Technology Acceptance Model (TAM)

The TAM constructs (PU, PEOU, and ITU) were used to evaluate the three proposed artifacts: i) the *SafeSecIoT Canvas* model, ii) the *STPA-SafeSecIoT* method, and iii) the *SafeSecRETS* tool. The constructs and items (questions) used in the evaluation questionnaires are presented below.

- **Perceived Usefulness (PU)**
  - It would allow me to complete tasks more quickly.
  - It would improve my performance at work.
  - It would increase my productivity at work.
  - It would improve my effectiveness at work.
  - It would make my work easier.
  - I would find it useful in my work.
- **Perceived Ease of Use (PEOU)**
  - Learning to use it would be easy for me.
  - I would find it easy to use it to do what I want.
  - My interaction with it would be clear and understandable.
  - It would be easy to interact with it.
  - It would be easy to become skilled at using it.
  - It would be easy to use.
- **Intention to Use (ITU)**
  - I would use this artifact.

  - I would recommend this artifact to academics or professionals.
  - I see value in adopting this artifact in the context of my work or research.

Task-Technology Fit (TTF)

In addition to TAM, a questionnaire based on TTF was applied to the tool:

- **Task-Technology Fit (TTF)**
  - It is suitable for supporting task completion.
  - It helps me perform tasks with greater accuracy and consistency.
  - It allows me to easily access the information and elements necessary for the task.
  - It correctly implements the features necessary for task completion.
  - It improves the quality and speed with which I perform tasks.

B.1.2 7-point Likert scale

For all questions presented above, a 7-point Likert scale was used to measure participants' responses. The following scale was adopted:

1. I strongly disagree
2. I disagree
3. I weakly disagree
4. I neither agree nor disagree
5. I weakly agree
6. I agree
7. I strongly agree

B.2 Overview of Results: TAM and TTF

B.2.1 *SafeSecIoT Canvas* Model: Detailed Responses

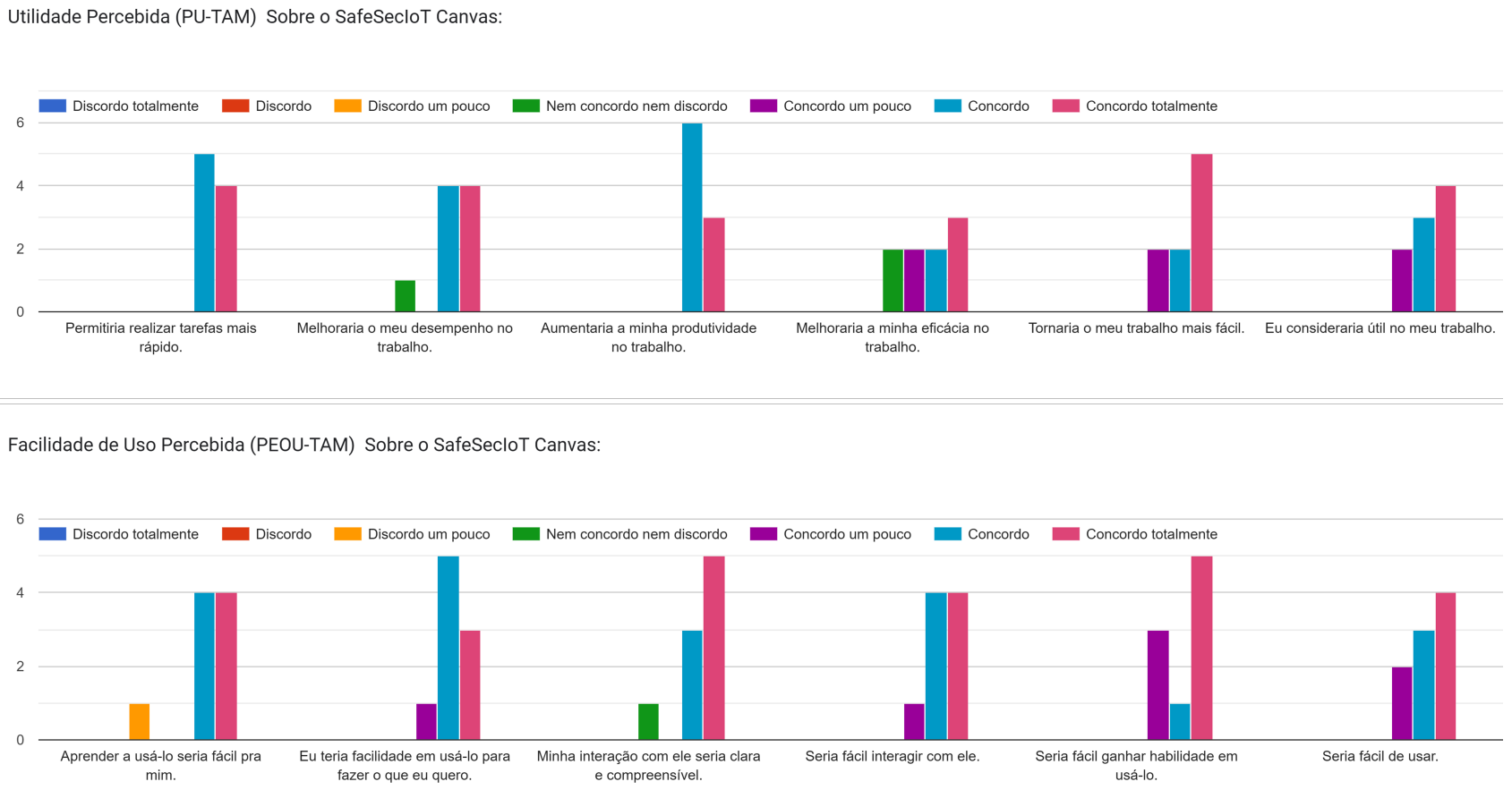


Figure B.1: Perceived usefulness and ease of use of the *SafeSecIoT Canvas* model.

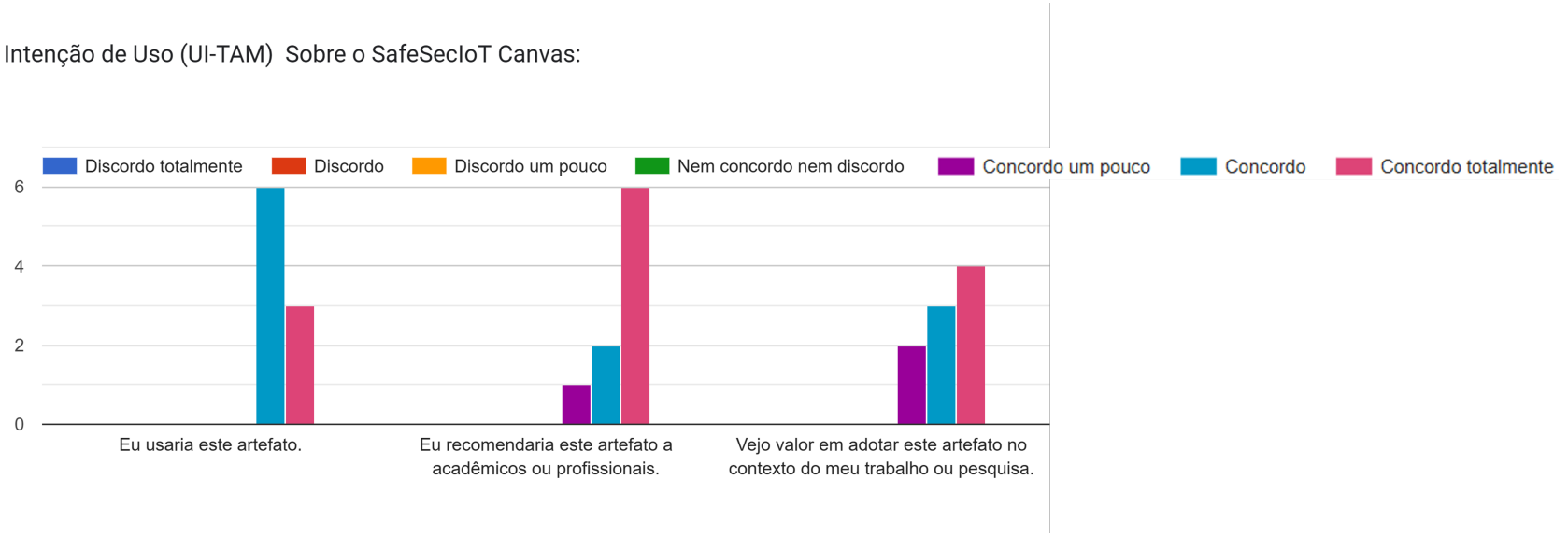
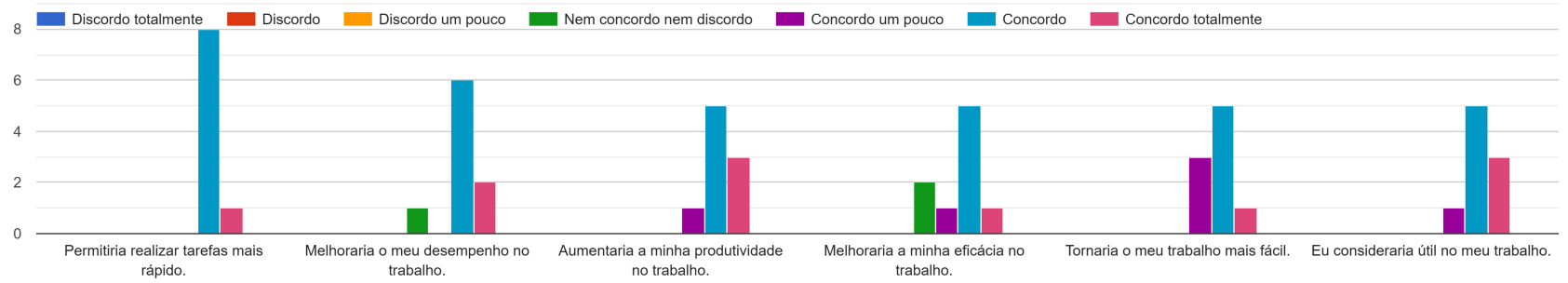


Figure B.2: Intention to use of the *SafeSecIoT Canvas* model.

B.2.2 *STPA-SafeSecIoT* Method: Detailed Responses

Utilidade Percebida (PU-TAM) Sobre o STPA-SafeSecIoT:



Facilidade de Uso Percebida (PEOU-TAM) Sobre o STPA-SafeSecIoT :

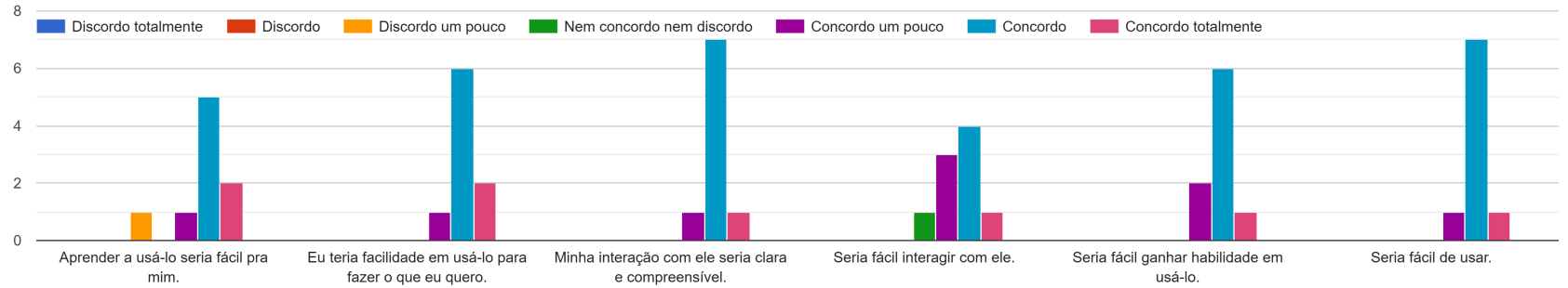


Figure B.3: PU and PEOU of the *STPA-SafeSecIoT* method.

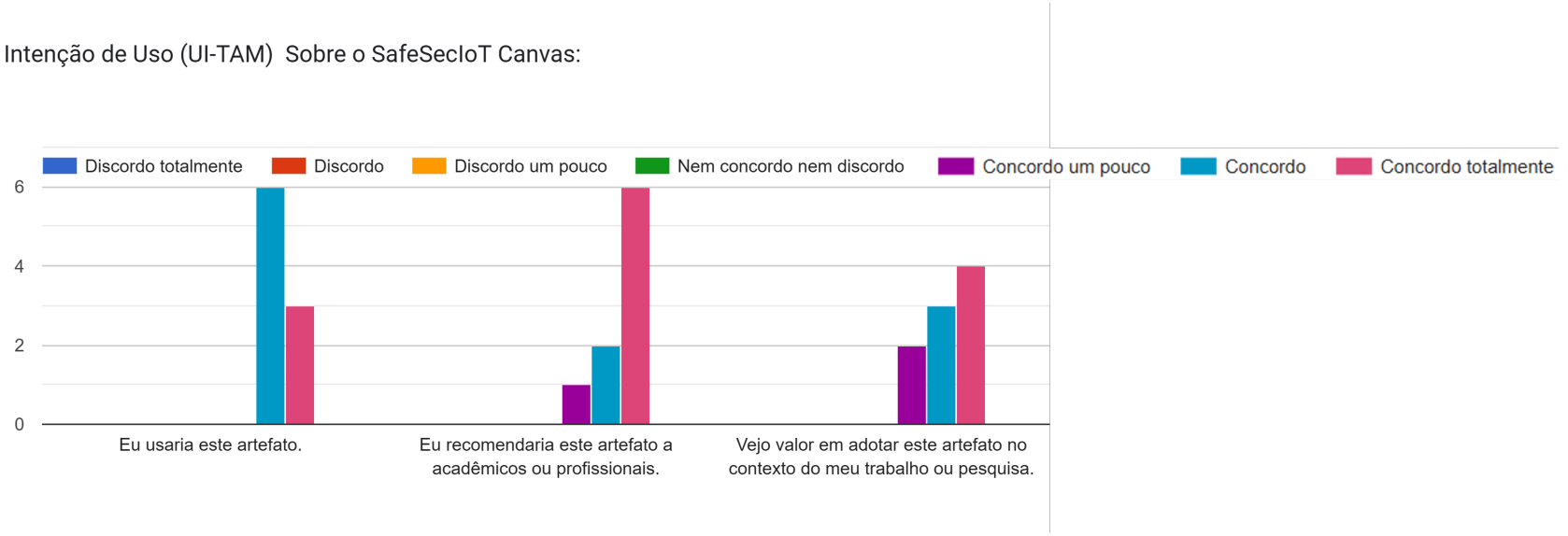


Figure B.4: Intention to use of the *STPA-SafeSecIoT* model.

B.2.3 *SafeSecRETS* Tool: Detailed Responses

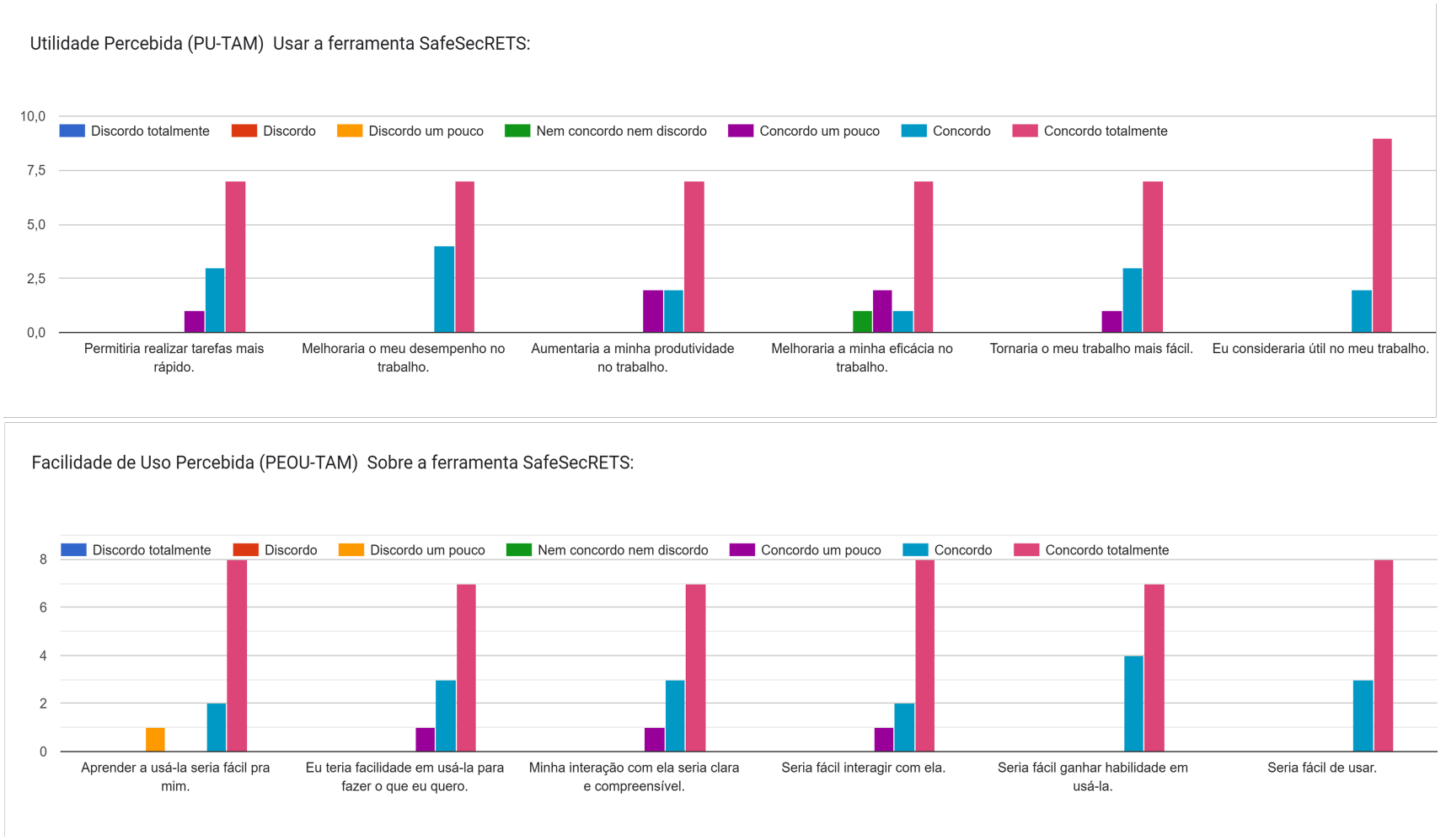


Figure B.5: PU and PEOU of the *SafeSecRETS* tool.

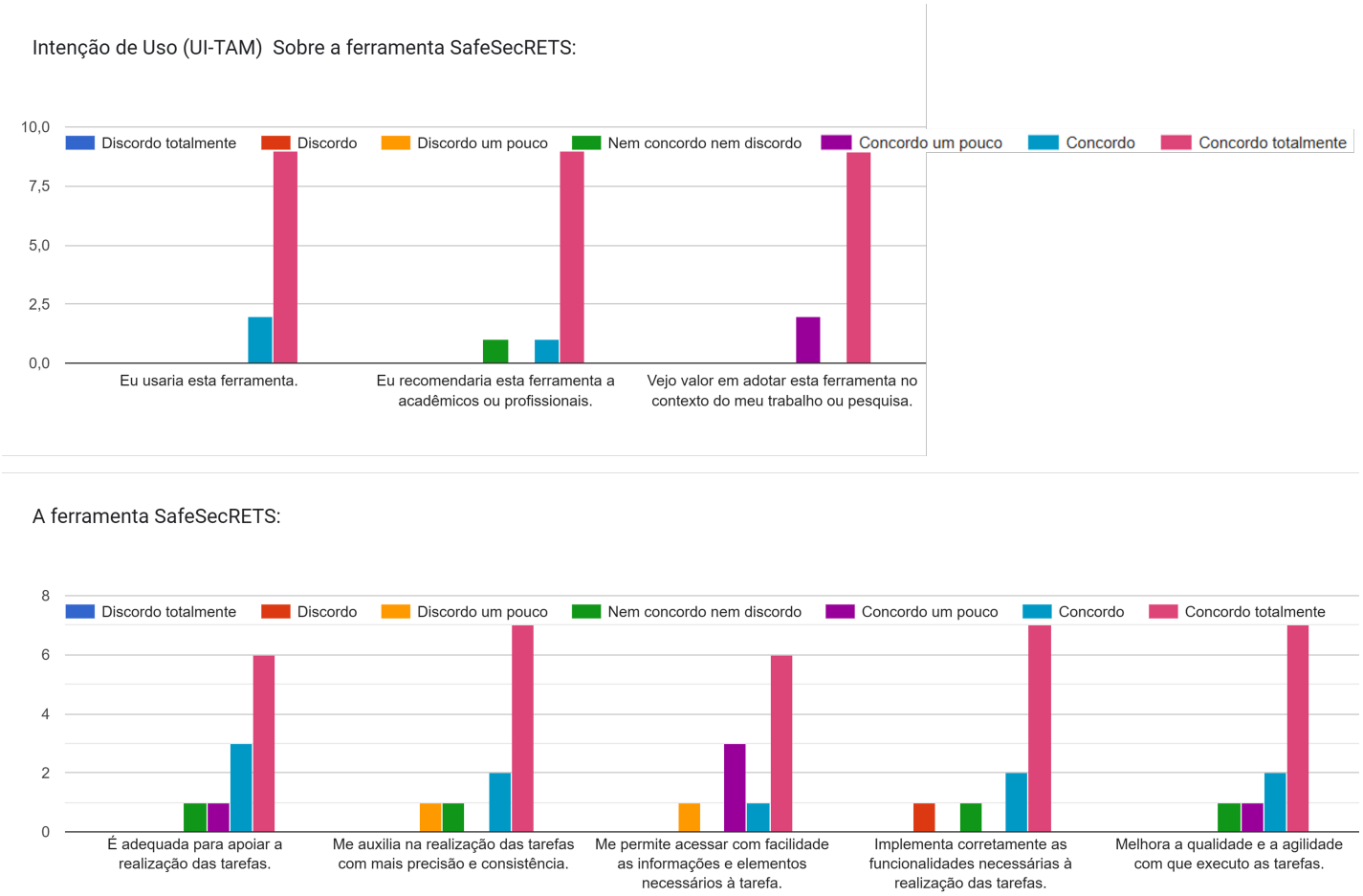


Figure B.6: ITU and TTF of the *SafeSecRETS* tool.

B.3 Support Material

In addition to the text of this thesis and published articles, we produced supporting materials for the evaluations carried out and also a website for the research. Below, we briefly present the other materials from this research with the respective links for access:

- The research website, containing supporting materials and links to the publications produced in this research: https://www.safesecrets.com.br/
- Videos presenting the proposal and the artifacts developed:
  - Video presenting the *SafeSecIoT Canvas* and *STPA-SafeSecIoT* artifacts: https://youtu.be/s8vlcSsUlwA
  - Video presenting the *SafeSecRETS* tool: https://youtu.be/WnTYjGzAwKs
- Link to the tool and test project: https://safesecrets.bubbleapps.io/version-test/
  - User: user@sbes.com
  - Password: 5azRdt7QTJ7VZE9