Lund University

Faculty of Engineering LTH

Department of Electrical and Information Technology

Master's Thesis

Privacy threat analysis and evaluation of Privacy Enhancing Technologies for the integration of 3D Scene Graphs in smart buildings

Alejandra Múgica Trápaga

September 2023

Master's Thesis

Title:	Privacy threat analysis and evaluation of Privacy Enhancing Technologies for the integration of 3D Scene Graphs in smart buildings
Department:	Department of Electrical and Information Technology
Author:	Alejandra Múgica Trápaga
Supervisor Ericsson:	Gregoire Phillips
Supervisor LTH:	Christian Gehrmann
Supervisor ETSIT:	Mario Vega Barbas
Examiner:	Thomas Johansson

Lund, of of 2023

Fdo.:

Abstract

The recent advent of 3DSG in the computer vision domain has brought powerful high-level representations of 3D environments. These representations strive to mimic human perception and facilitate the extraction of meaningful insight from visual data. This research has led to framework like Hydra, which creates a real-time, persistent spatial perception system that creates a complete 3D map of their surroundings. Furthermore, smart cities are at the forefront of integrating information technology into urban life, with smart building playing an essential role. Smart buildings are equipped with IoT devices that promote their interaction to achieve a more intelligent environment. The integration of spatial perception systems like Hydra into the smart building domain can improve decision-making, autonomy, and responsiveness. However, this integration introduces a new dimension of privacy concerns among users. This arises due to Hydra's capacity to capture sensitive information, raising valid concerns regarding the potential misuse by third parties or attackers. This necessitates a comprehensive approach that helps practitioners to systematically identify and assess potential threats and vulnerabilities to a system, application, or network. This thesis places a primary focus on addressing privacy concerns within the integration of Hydra into the smart building domain. To achieve this, this study applies the LINDDUN privacy threat methodology, which is designed to focus on privacy concerns, guiding practitioners in identifying and mitigating privacy threats. Thus, the main the goal of this thesis is to thoroughly examine and address these threats, all with the ultimate goal of protecting personal information.

Keywords: 3DSG, DFD, Hydra, IoT, LINDDUN, PETs, Privacy, Threat modeling.

Contents

List of Acronyms

3DSG	**3D** **S**cene **G**raph
AR	**A**ugmented **R**eality
DFD	**D**ata **F**flow **D**iagram
DOF	**D**egrees **of** **F**reedom
GDPR	**G**eneral **D**ata **P**rotection **R**egulation
IEC	**I**nternational **E**lectrotechnical **C**omission
IMU	**I**nertial **M**easurement **U**nit
IoI	**I**tem **o**f **I**nterest
IoT	**I**nternet **o**f **T**hings
ISO	**I**nternational **S**tandarization **O**rganization
LINDDUN	**L**inkability **I**dentifiability **N**on-repudiation **D**etectability **D**isclosure of information **N**on-compliance
MR	**M**ixed **R**eality
NIST	**N**ational **I**nstitute **S**tandars and **T**echnology
OCTAVE	**O**perational **C**ritical **T**hreat, **A**sset, and **V**ulnerability **E**valuation
ORRM	**O**WASP **R**isk **R**ating **M**ethodology
OWASP	**O**pen **W**eb **A**pplication **S**ecurity **P**roject
PETs	**P**rivacy **E**nhancing **T**echnologies
PII	**P**ersonal **I**dentifiable **I**nformation
QoS	**Q**uality **o**f **S**ervice
RL	**R**einforcement **L**earning
RGB-D	**R**ed, **G**reen, **B**lue, **D**epth
SfM	**S**tructure **f**rom **M**otion
SDL	**S**ecure **D**evelopment **L**ifecycle
STRIDE	**S**poofing **T**ampering **R**epudiation **I**nformation **D**isclosure **D**enial of **S**ervice **E**levation of **P**rivilege
VAST	**V**isual **A**gile **S**imple **T**hreat

1. Introduction, objectives and thesis outline

1.1 Introduction

The recent advent of 3D Scene Graphs (3DSG) has brought powerful high-level representations of 3D environments. 3DSG serve as an objective semantic representation of the scene that effectively captures the underlying structure, relationships and context allowing us to obtain valuable insights from visual data. These representations aim to resemble how humans naturally perceive and interpret visual information. Humans perceive the three-dimensional structure of objects and scenes with ease, while computer vision relies on mathematical techniques and algorithms to reconstruct 3D shape and appearance from images. Acquiring these representations stands as a central focus on researchers' investigations, leading to frameworks like Hydra [1]. This innovative framework describes a real-time and persistent Spatial Perception System that allows building a 3D map of the surroundings, representing objects, their semantic labels and the relationship among these entities. Hydra addresses both the perception and planning aspects of map representations in the robotics field. They involve the creation of high-quality 3D reconstructions of natural environments from sensor data and the creation of maps to navigate through the environment in a safe and collision-free manner. For these purposes, Hydra enables the extraction of relevant information from the 3D map and to query the location of objects, rooms, or moving people.

Furthermore, smart cities strive to integrate information technology into every aspect of city life, aiming to improve their citizens quality of life and to create economic growth. One of the key areas of smart cities is smart buildings, which seek to create a more efficient environment to work or live. Smart buildings are becoming equipped with devices that can communicate and interact with each other, often referred to as the Internet of Things (IoT). The integration of frameworks like Hydra into the smart building scenario holds significant potential for enhancing the capabilities of connected devices. By leveraging Hydra's advanced perception and planning functionalities, IoT devices can achieve more robust and efficient interactions with their surroundings. This can lead to improved decision-making, autonomous navigation and intelligent responses. 3DSG opens the way for smarter and more context aware IoT devices that can better understand, interpret and interact with the physical world. Moreover, the incorporation of frameworks like Hydra into the smart building domain introduces a new dimension, giving rise to privacy concerns among users. Perceptual applications such as Hydra pose an inherent risk of capturing sensitive information and potentially misused by third parties or attackers and require careful consideration.

Privacy has continually been shaped by the rapid advancements in technology and it is seen as the right to have control over how personal information is collected, processed and stored. The growing importance of data privacy in the digital age has led to the implementation of regulatory measures aimed at protecting individuals' personal data. For this purpose, the European Parliament approved the General Data Protection Regulation (GDPR) in 2018 [2]. This comprehensive privacy and security data protection law require organizations to follow its principles and helps them during the design phase. Moreover, organization like the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have developed several international standards addressing privacy concerns [3]. These laws and standards require organizations to incorporate data protection of personal identifiable information (PII) as a fundamental principle during the design of their activities, ensuring a transparent data usage and implementing data protection techniques.

Additionally, protecting PII requires a comprehensive approach that help practitioners to systematically identify and assess potential threats and vulnerabilities to a system, application, or network. Threat modeling is a proactive process that aids in recognizing and evaluating potential threats that can negatively impact a particular entity. The LINDDUN privacy threat modeling framework was defined by KU Leuven in 2010 and a recent study in 2022 has acknowledged this approach as one of the two most recognized methodologies and among the ten most frequently used methodologies in academic literature [4]. Standards like the ISO TR 27550 on Privacy engineering for system life cycle processes [5], and organizations such as the National Institute of Standards and Technology (NIST) [6], among others, have acknowledged this methodology for conducting privacy threat analysis. This methodology is centered around seven privacy threat categories: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness and Non-compliance. It follows a structured sequence of steps that guide practitioners in identifying and mitigating potential threats by following the distinct threat categories and proposing privacy-enhancing technologies (PETs). Ultimately, the outcome of this process is a comprehensive privacy risk assessment that provides insights into the potential privacy threats and recommendations for implementing PETs to protect PII, while upholding privacy standards and regulations [7].

1.2 Research question and objectives

The main research question of this thesis explores the implementation of the LINDDUN privacy threat model to 3DSG in smart building scenarios. Specifically, this research focusses on a spatial perception system known as Hydra. This research aims to highlight the critical but often neglected necessity to conduct a comprehensive privacy threat analysis within this domain. The protection of individuals' privacy is of the utmost concern in this context, as it involves the potential collection and possible misuse of sensitive and personal information. Therefore, this study places a strong emphasis on addressing these privacy concerns as a key research focus.

Following the stated research question, this project is guided by two main objectives. The first objective is to conduct a privacy threat identification. This involves systematically cataloging potential threats and vulnerabilities that could compromise individuals' privacy when integrating a framework like Hydra into the smart building scenario. This step includes an examination of how sensitive and personal information might be captured, accessed, or potentially misused. It also includes identifying vulnerabilities during data transmission, storage and processing. The goal is to create a comprehensive inventory of these threats, ensuring a thorough examination and documentation of all privacy-related threats. To achieve this objective, I will present two different scenarios in which Hydra assumes distinct roles. The first scenario involves examining Hydra as a subcomponent integrated into the context of a smart building. In this scenario, Hydra interact with other technologies and entities within the smart building ecosystem. The second scenario focuses on Hydra as a unique process. In this case, I will thoroughly explore the subprocesses that form Hydra to understand its working in more detail. The second objective is centered on proposing mitigation strategies aimed at ensuring the protection of individuals' privacy. This task involves proposing PETs and countermeasures to mitigate the identified threats, with its main goal being the protection of sensitive information at any stage of the process.

1.3 Thesis outline

The outline of the report is structured as follows: Chapter 2 includes a literature review, focusing on the key aspects pertinent to this thesis. The primary subjects under consideration include Hydra's as a real-time 3DSG generation within the computer vision field, the significance of privacy regulations and standards mandating privacy assessment during the design phase and the role of smart buildings in the broader context of smart cities and IoT technology. Chapter 3 offers a thorough overview of the LINDDUN privacy threat methodology. It begins by describing the various privacy threat categories and outlining the necessary steps for a comprehensive privacy assessment. These steps, which serve as a guide in the following chapters, are organized into two main categories; the problem space, aimed at achieving a detailed documentation of the potential privacy threat scenarios and the solution space, which centers on offering a systematic proposal of PETs. Chapter 4 focuses on the problem space of the LINDDUN methodology, with the goal of documenting potential privacy threats. This aligns with the first main objective of this thesis. The problem space involves three steps: defining the data flow diagram (DFD), mapping the elements of the DFD to the privacy threat categories and identifying potential privacy threats. Chapter 5 focuses on the solution space of the LINDDUN methodology, with the goal of proposing suitable PETs for the identified threats. This aligns with the second main objective of the thesis. The solution space involves three steps: prioritizing threats, eliciting mitigations strategies and selecting corresponding PETs. Chapter 6 serves as the concluding section of this thesis, providing a summary of the key insights generated throughout the study. Within this chapter, the primary emphasis is placed upon deriving conclusive findings from the research and addressing the research question. Furthermore, this chapter introduces potential directions for future research within this field of study.

2. Literature Review

In this chapter, a comprehensive literature review is conducted to present the main theoretical foundations related to privacy, computer vision and smart buildings.

2.1 Privacy

Throughout the ages, individuals have highly regarded privacy and the protection of personal information. The debate on privacy in the western world dates to the introduction of the newspaper printing press and photography, which prompted Samuel D. Warren and Louis Brandeis to write their influential article on privacy in the Harvard Law Review in 1890. They argued for a “right to be left alone” and emphasized the principle of “inviolate personality” as the response to intrusive journalistic practices. Since then, the conversation surrounding privacy has evolved, emphasizing the individual’s right to control the extent of access to their personal information by others, while also acknowledging society’s right to access information about individuals [8].

The right to privacy was established in the 1950 European Convention on Human Rights, which affirms the importance of respecting one’s private and family life, home and correspondence. Since then, the European Union (EU) has regarded the right to privacy as a fundamental starting point and has ratified laws and regulations to protect and guarantee this right. The EU’s first step in this direction was the establishment of the European Data Protection Directive in 1995, which laid down the minimum data privacy and security standards. The rapid advancement of technologies and the increasing prevalence of the Internet created a need for an updated and data-centered approach to privacy standards. Consequently, in 2018, the General Data Protection Regulation (GDPR) was enacted after approval from the European Parliament [2]. Additionally, institutions like the International Organization for Standardization (ISO) that develop international standards for a wide range of activities have focused on privacy concerns. In these standards, experts provide optimal guidelines to approach tasks in a manner that upholds privacy principles [3].

Moreover, privacy debates have continually been shaped by the rapid advancements in technology and this association between privacy and access to information has deepened with the progress of information technology. The ever-expanding capabilities of digital systems and the pervasive nature of data-driven technologies add new dimension to discussions surrounding privacy, exploring the delicate balance between safeguarding personal information and harnessing the power of data. Technology’s continuous advancement leads to additional risks to individuals’ privacy, prompting an increasing demand for robust privacy frameworks and stronger data protection laws [8].

2.1.1 Personal data and moral reasons for protecting it

PII refers to any data that attackers can use to potentially identify a specific individual, whether directly or indirectly. It includes data such as names, social security numbers, birthdates and even details like medical records, education, finances and employment information that can be linked to an individual [9]. Moreover, the GPDR defines personal data as “any information that relates to an individual who can be directly or indirectly identified”. The notion of this definition encourages a comprehensive understanding of personal data, encompassing various types of information that can be used to identify individuals. This includes explicitly provided details like a person’s name, identification number, date of birth, location data or address as well as indirect information that can reveal their physical, physiological, genetic, mental, commercial, cultural, or social identity. Furthermore, the scope of personal information extends to less explicit information, including work time recordings that indicate an employee’s start and end times, breaks and non-work hours, or subjective information like opinions or judgements [10].

The protection of personal data and the control over access to such data are driven by several moral reasons. First, unrestricted access to personal data, such as bank accounts or social media accounts, can lead to potential harm for the data subject in diverse ways. Second, personal data have become commodities, placing individuals in an unfavorable position to negotiate contracts regarding their data usage and lacking the means to verify if partners uphold the contract terms. Data protection laws aim to establish fair conditions for personal data transmission and exchange, mitigating the issues related to asymmetrical information. Third, discrimination and disadvantages for individuals may arise when personal information from one field, like health care, undergoes a shift in meaning when used in another field or context, like commercial transactions. Fourth, mass surveillance and the lack of privacy can limit the independence and personal autonomy, undermining individuals’ sense of freedom to act independently and their dignity. Reducing individuals to mere data points neglects their unique qualities and moral agency [8].

Data protection applies solely to information about a natural individual, starting from their legal capacity at birth until their death. Article 6 of the GDPR outlines the legal grounds for processing personal data. Without a valid justification, any collection, storage, or sale of personal data is prohibited. Acceptable reasons for processing include obtaining specific consent from the data subject, executing or preparing a contract involving the data subject, complying with a legal obligation, saving someone’s life, performing a task in the public interest or an official function, or having a legitimate interest in processing personal data. However, even with a legitimate interest the rights and freedoms of the data subject always take precedence, particularly when it concerns children’s data. Once you determine the lawful basis, it becomes essential to document it and notify the data subject transparently. Any changes in the justification must be well-founded, documented and communicated to the data subject for transparency [11].

2.1.2 Privacy by design and by default

Organizations must consider the data protection principles of privacy by design and privacy by default as fundamental pillars of their operations to ensure that privacy is integrated into every aspect of their products, services and systems. In a general sense, Privacy by Design (PbD) is a comprehensive concept that involves various practical elements. PbD states that practitioners must proactively address privacy concerns by addressing mitigation of privacy concerns early into the development cycle and involving qualified expertise to guide the process. Moreover, experts need to adopt and integrate Privacy-Enhancing Technologies (PETs) with the main purpose of respecting and protecting users’ privacy.

This approach has been endorsed by Data Protection Authorities, legally mandated by GDPR and supported by the European Commission to foster the data economy [12].

Article 25 of the GDPR emphasizes the importance of data protection by design and by default. It states that controllers must implement suitable technical and organizational measures to protect data subjects' rights and comply with the regulation's requirements. With this intention, controllers must consider various factors, including the latest advancements in technology, implementation costs and the nature, scope, context and purposes of data processing, along with the potential risks to individuals' rights and freedoms resulting from the processing. Additionally, the controller must ensure that, by default, only necessary personal data are processed for each specific purpose, limiting the amount, extent, storage period and accessibility of the data. An approved certification mechanism can be used to demonstrate compliance with these requirements, in accordance with Article 42. This article of the GDPR promotes the establishment of data protection certification mechanisms, seals and marks to demonstrate compliance with the regulation for data processing by controllers and processors [13].

2.1.3 General Data Protection Regulation

The General Data Protection Regulation (GDPR) is a comprehensive and strict privacy and security data protection law that was introduced by the EU on May 25, 2018, replacing the Data Protection Directive 95/46/EC. The main purpose of the GDPR is to strengthen and unify data protection for individuals, ensuring a robust and unified approach to safeguard personal data. It applies to any organization that targets and processes personal data of EU residents, regardless of whether the organization is based in the EU or not. The GDPR protects all EU individuals and grants them more control over their personal data when shared with organizations [2].

This regulation defines various participants involved in the data processing, encompassing any action performed on data, whether automated or manual. Firstly, the term data subject refers to the individual whose data is being processed, usually visitors or customers and can be considered as the owner of the data. The data controller, typically an owner or employee within an organization, holds the responsibility for determining why and how personal data will be processed. Lastly, data processors are third parties that handle personal data on behalf of a data controller. In line with the GDPR's mandate to protect individuals' data and privacy, data subjects gain several privacy rights. These rights are interlinked with diverse articles within the GDPR, which serve as a legal framework to enforce and protect these rights. The relation between the privacy right and the GPDR articles is outlined below [2].

1. The right to be informed: This encompasses Article 12, which stipulates the need for transparent information, communication and clear methods for exercising data subject rights. Additionally, Article 13 mandates the provision of information when personal data is collected from data subjects and Article 14 covers information dissemination when data is sourced from entities other than the data subject.

2. The right of access: Article 15 grants individuals the right to access their personal data, providing transparency and control over their information.

3. The right to rectification: Under Article 16, individuals possess the right to rectify inaccurate or incomplete personal data.

4. The right to erasure: Commonly known as the ”right to be forgotten,” Article 17 empowers individuals to request the removal of their personal data under specific circumstances.
5. The right to restrict processing: Article 18 outlines the right to request the restriction of data processing in certain situations.
6. The right to data portability: Article 20 grants individuals the ability to receive their personal data in a structured, commonly used and machine-readable format, enhancing data mobility.
7. The right to object: Article 21 provides individuals the right to object to certain types of data processing.
8. Rights in relation to automated decision making and profiling: Article 22 delves into provisions concerning automated decision-making and profiling, safeguarding individuals from potential adverse effects.

Furthermore, the GDPR requires data subjects to explicitly consent to the processing of their personal data. Article 7 of the GDPR outlines the strict rules that controllers must follow to obtain consent and demonstrate that consent has been given by the data subjects [14]. These rules state that consent must meet specific criteria, including being “freely given, specific, informed and unambiguous” and data subjects have the right to withdraw consent at any time. For organizations, consent requests should be easily distinguishable from other matters and presented in a clear and straightforward manner. Once consent is obtained, organizations are prohibited from changing the legal basis of processing to another justification. During this process, it is indispensable to keep documentary evidence for compliance purposes. On another note, the GDPR outlines seven essential regulatory aspects, described in the Article 5.1-2, that must be adhered to when dealing with consent and data processing [15].

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitation
- Integrity and confidentiality
- Accountability

These guidelines aim to ensure that organizations process data in a lawful, fair and transparent manner while respecting individuals’ privacy rights. These principles include processing data only for specified legitimate purposes, collecting and processing the minimum data, maintaining data accuracy, storing data for the required duration and implementing appropriate measures to ensure data integrity and confidentiality. By adhering to these principles, data controllers have to be able to demonstrate accountability and responsibility in their data processing practices. For this purpose, organizations designate a data protection responsible and, in some cases, appoint a Data Protection Officer (DPO) to maintain detailed documentation of the data that is being collected, processed and stored [15].

2.1.4 ISO Standards

ISO/IEC 27000 family

The ISO/IEC 27000 family of standards [16] offers an overview of Information Security Management Systems (ISMS) that organizations can integrate into its regular practices. An ISMS encompasses policies, procedures, guidelines and resources with the main objective of securing information aligning with business goals. Organizations, regardless of their size and type, engage in the collection, processing, storage and transmission of information. Organizations must understand the significance of information and related elements as important assets for achieving their objective, while also acknowledging the potential risks that could impact these assets. Thus, information is an important asset in today's organizational operations that must be adequately protected. These standards address the importance of organization implementing measures to address these risks and taking steps to address these risks though information security controls. They must regularly oversee the current procedures to identify emerging risks and select, implement and update controls as necessary. Thus, the aggrupation of standards under the USI/IEC family recommends methods for effectively managing information risks through the implementation of security controls. The relations between these standards are shown in Figure 2.1 [16][17] .

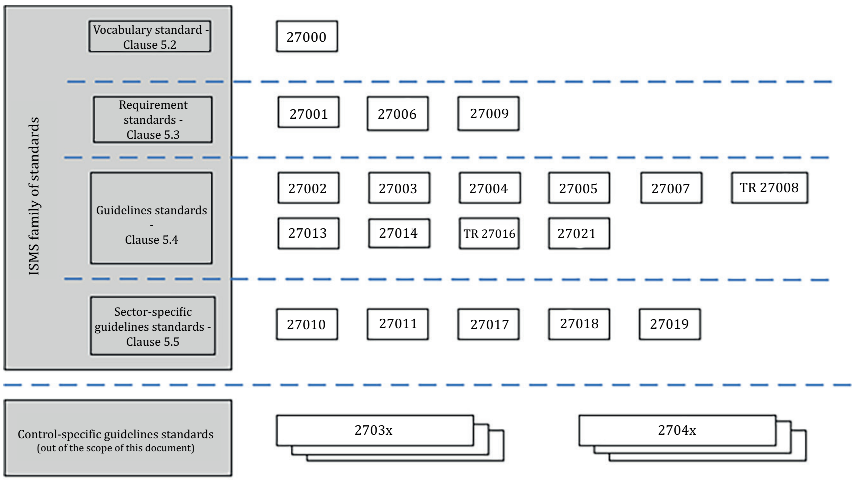

ISMS family of standards
Vocabulary standard - Clause 5.2
27000
Requirement standards - Clause 5.3
27001
27006
27009
Guidelines standards - Clause 5.4
27002
27003
27004
27005
27007
TR 27008
27013
27014
TR 27016
27021
Sector-specific guidelines standards - Clause 5.5
27010
27011
27017
27018
27019
Control-specific guidelines standards (out of the scope of this document)
2703x
2704x

Figure 2.1: ISMS family of standards relationships [16].

These inter-related standards focus on several structural elements, such as standards outlining ISMS requirements (ISO/IEC 27001), certification body requirements (ISO/IEC 27006) for certifying conformity with ISO/IEC 27001 and a requirement framework for sector-specific ISMS implementations (ISO/IEC 27009). Additionally, there are guidance documents that cover different aspects of ISMS implementation, encompassing both a general process and specific guidance for different sectors. By implementing these standards, organizations can reduce information security risks, decreasing the likelihood and impact of security incidents [16].

ISO 29100

The ISO 29100 [18] is an international standard that establishes a comprehensive framework for safeguarding PII in Information and Communication Technology systems (ICT). This privacy framework addresses the ever-growing challenges posed by the commercial utilization and significance of PII, the cross-border sharing of PII and the escalating complexity of ICT systems. It provides organizations with a set of guidelines and principles to manage and mitigate risks associated with PII. This standard was last amended in 2018 and the second edition is now being under development.

Similar to the GDPR, this standard defines several participants involved in the data processing and the interactions between these actors. The PII principals are the PII owners and they share it with the PII controllers and the PII processors for processing, requiring their consent. PII controllers determines the purpose and means of why and how the PII is processed. This actor is responsible for upholding to the privacy principles during processing. The PII processors are the actors responsible for conducting the processing on behalf of the PII controllers, following the instruction provided by the PII controller. Furthermore, either a PII controller or a PII processor can transfer PII to a third party. Third parties do not handle PII on behalf of the PII controller and, once they have received the specific PII, this actor assumes the role of the PII controller in its own capacity. All of the actors possess the potential to engage in diverse interactions with one another, encompassing a range of possibilities for potential pathways of exchange of information.

This standard defines privacy requirements that organizations must follow in order to protect PII during its processing. These requirements are part of the risk management process, which is shaped by legal and regulatory mandates, contractual considerations, specific business needs among other factors. Additionally, this standard emphasizes the need for organizations to adopt a suitable privacy policy and privacy controls for the handling of PII. Privacy policies should align with the organization's purpose, outline objectives, commit to privacy requirements and ensure continuous improvement and to be communicated internally and externally. On the other hand, organization should adopt privacy controls based on the requirements from the risk assessment. It highlights the integration of privacy controls into the design phase as part of the "privacy by design" approach and tailoring information security controls based on specific PII processing risks. Furthermore, this standard outlines several privacy principles, which are listed below and act as guiding directives in structuring the design, development and implementation of privacy policies [18].

- Consent and choice
- Purpose legitimacy and specification
- Collection limitation
- Data minimization
- Use, retention and disclosure limitation
- Accuracy and quality
- Openness, transparency and notice
- Individual participation and access
- Accountability
- Information security
- Privacy compliance

2.2 Computer vision

Computer vision is an Artificial Intelligence (AI) field that enables machines to observe and understand visual information from their surroundings, extracting meaningful insights from digital images and videos. This process aims to resemble how humans perceive visual information. However, unlike the effortless way in which humans interpret their environment, computer vision algorithms face significant challenges and are prone to errors. Humans naturally perceive the three-dimensional structure of objects and scenes with ease, while computer vision relies on mathematical techniques and algorithms to reconstruct 3D shape and appearance from images. Researchers have made significant progress in computer vision, including accurate 3D modeling, object tracking and facial recognition. However, achieving image interpretation at the level of a human remains challenging due to the complexity of modeling the visual world. Computer vision trains machines to identify objects, distances and movements, while also recognizing patterns, shapes, colors and textures. Additionally, it can identify and track objects, detect and classify human faces and extract useful information from visual data. The industries leveraging from computer vision range from the energy and utilities to manufacturing and automotive and it is a rapidly evolving field that is expected to continue expanding [19][20].

Furthermore, any discussion of map representations in robotics generally has two sides: the perception side, which often focuses on creating high-quality 3D reconstructions of natural environments from sensor data and the planning side, which uses pre-built maps to navigate through the environment in a safe and collision-free manner. Mapping and planning often have very different requirements from an environmental representation. From the perception standpoint, it is often most important to be able to output a high-quality coloured surface model, such as a mesh for mapping. On the other hand, for navigation and planning, it is often most essential to have fast collision checking and be able to compute clearance and direction toward nearest obstacles. Therefore, high-level representations are needed to comprehend and carry out human instructions, while also facilitating fast planning or mapping processes [21].

2.2.1 Levels of perception

The cognitive process relies heavily on the close relation between perception and cognition. On one side, perception is the process of recognizing, organizing and interpreting sensory information from the environment through our senses, which helps us understand and make sense of the world around us. On the other hand, cognitive science investigates the nature of the mind, how it works and the way in which humas perceive the world. This field of study brings together different areas like psychology, neuroscience, computer science, linguistics and philosophy. The main focus of this field is trying to understand how human minds consistently process raw amounts of data from their environments, adeptly discerning patterns, extracting meaning and identifying significance within the intricate complexity of information. Decision making and problem solving are two concepts that cognitive science tries to replicate by creating models or systems that simulate the way human minds navigate choices and find solutions. In summary, perception enables a conscious engagement with the surrounding environment, while cognition empowers to develop beliefs, reach decision and more [22][23].

The notion of cognition has evolved since Kant's statement "Concepts without percepts are empty; Percepts without concepts are blind" [23]. This highlights the interdependency of perception and cognition. Kant divided the perceptual work of the mind into the faculty of sensitivity and the faculty of understanding. Sensibility is responsible for gathering raw sensor information, while understanding structures this information into a coherent and meaningful perception of the world. Although this statement is long due, the concept of perception remains relevant. In line with Kant's sensibility faculty, low-level perception covers the initial sensory data reception, while higher-level perception involves a broader perspective. It grasps the overall significance and coherence of a situation in a more profound and conceptual manner. Thus, high-level perception begins when thinking and understanding becomes important in how we make sense of things we see or experience, involving more complex cognitive elements. One key feature is that individuals can understand the same information in many different ways, which makes high-level perception very flexible. This difference in understanding depends on the context and the state of the individual, which can be influenced by beliefs, goals, or external context. This results in a variety of understanding of the same environment, which indicates the importance of semantics in representation. Furthermore, high-level perception is closely related to the problem of mental representation. Representations result from the process of shaping raw data into organized and coherent structures, which the mind utilizes for various purposes. Research in the field of AI have extensively been exploring the process of representations, including challenges of determining relevance and organizing data for effective and meaningful representation [23].

2.2.2 Scene Graphs and 3D Scene Graphs

The primary objective of computer vision is to achieve a comprehensive understanding of visual scenes, extracting valuable insights from the input visuals. Researchers strive for visual scene understanding that seek to emulate the process of how humans interpret visual information. These objectives represent significant and challenging tasks within the field of computer vision. Visual scene understanding goes beyond the tasks of recognition and localization of objects and aim to address higher-level challenges that involve semantic relationships between objects and their interactions with the environment. For this purpose, researchers seek to build efficient structured representations that capture semantic understanding of the visual scenes [24].

A scene graph is a structural representation that captures visual information from a particular scene along with detailed semantics. These representations aim to provide a clear and unbiased representation of the elements, attributes and interconnections within a scene by accurately capturing the content and relationships present. In scene graphs, nodes correspond to object instances accompanied by their attributes, while edge symbolize the relationships that exist between these instances. Object attributes range from physical features such as color or material to dynamic aspects like their state. Relations can include a variety of different actions, spatial position, descriptive verbs, prepositions, or comparatives. Therefore, a scene graph consists of a set of visual relationships triplets as <*subject, relation, object*>as shown in Figure 2.2 [24].

This graph serves as an objective and unbiased representation that connects visual and semantic understanding, leading to a comprehensive understanding of the scene. Scene Graph Generation (SGG) involves the automatic mapping of images or videos into semantic structural scene graphs. To extend this concept to 3D space, researchers are working on

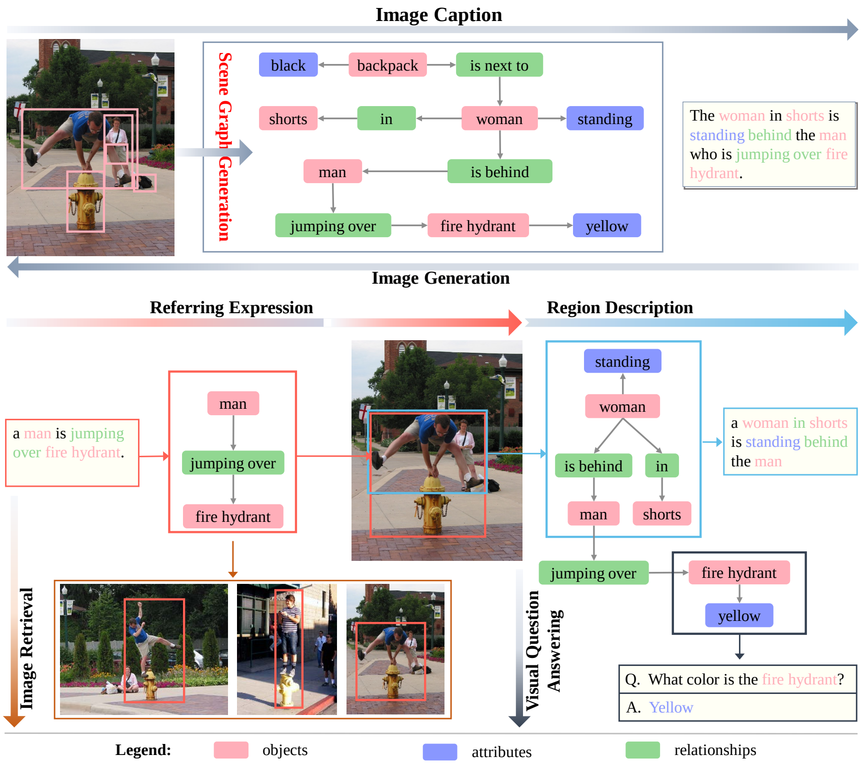


Figure 2.2: Visual illustration of a scene graph structure and some applications [24].

creating a structured format that can effectively capture and represent 3D information. In the 3D domain, there are various approaches for three-dimensional representations like multiple views, point clouds, polygonal meshes, wireframe meshes and voxels. Multiple views involve capturing the scene from different angles or viewpoints to create a sense of depth and dimensionality. Point clouds involve collecting a set of 3D points that collectively define the shape and structure of an object, where each point represents a specific coordinate in space. Polygonal meshes collectively create a surface that represents the object's shape by connecting vertices with edges to form polygons. Wireframe meshes represent objects using lines and edges to outline their form. Voxels, which stands for "volumetric pixels", divide the space into small cubic-cells that represent objects as a grid of 3D pixels, analogous to how 2D images consists of pixels. Thus, 3D scene graphs (3DSG) are high-level representation that enable the rendering of three-dimensional spaces as a layered graph [24].

2.2.3 Hydra

Hydra is an innovative 3DSG that generates a real-time and persistent representation of the environment, facilitating just in time decision making and ensuring long-term autonomy. This framework captures the metric and semantic aspects of the scene and is adaptable to expansive environments. It is built from sensor data, which is incrementally integrated into the 3DSG while the sensor is actively exploring its surroundings. To achieve this, it relies on odometry to estimate the robots' trajectory and relative position. Odometry continually estimates the change in position by using visual-inertial odometry (VIO), which is the process of estimating the pose and velocity of an agent. Hydra gradually depicts the scene as a layered graph at different levels of abstraction, where nodes describe object instances and edges relations between these instances, as shown in Figure 2.3 [1].

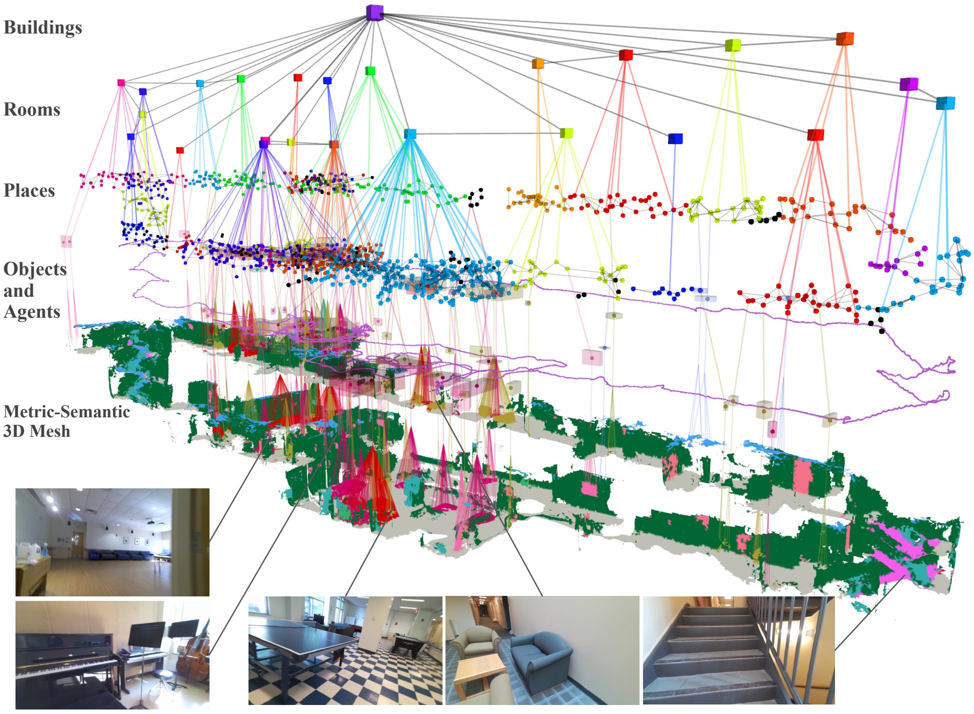


Figure 2.3: Generation of Hydra as a layered graph [1].

This example depicts an indoor environment with five different layers, in which nodes can go from low-level geometry to high-level semantics, including objects, agents, places, rooms and buildings. Firstly, the metric-semantic mesh layer contains all the nodes of static and dynamic elements. These nodes are linked to the object nodes they belong to in the following layer. Secondly, the objects and agents layer contain the object nodes, disregarding the structural elements and the nodes that belong to the time-varying entities. The objects and agents from the second layer are linked to the nearest place node. Thirdly, the places layer describes the free space paths. Fourthly, the rooms layer groups the places that belong to the same room. Lastly, a single building node groups all of rooms' nodes. Moreover, the construction of these layers is directed by the following approaches.

Metric-Semantic Mesh

This layer forms a real-time a metric-semantic 3D mesh by extending the work of Kimera [25]. Hydra modifies the module Kimera-Semantics principally based on Voxblox [26] and the marching cubes algorithm to form a volumetric model of the surroundings in a predefined radius.

Voxblox is a system designed for 3D modeling and mapping of the environment, enabling robots or other devices to navigate and interact with the environment in real-time. It can also be used for tasks such as object recognition, tracking and localization using sensors like RGB-D cameras. This tool uses a voxel-based approach to represent the environment, where the space is divided into voxels and its grouping is called voxel grids. Each voxel stores information about the occupancy, color and other properties of the environment at that location. These voxels are grouped into voxel grids and according to the information they contain, they are classified into Euclidean Signed Distance Fields (ESDFs) or Truncated Signed Distance Fields (TSDFs). While both of them are signed distance fields (SDFs), they differ in the way that the distance of a voxel is computed. ESDFs are voxel grids containing the Euclidean distance to the nearest occupied voxel and are usually used for collision checking, inferring distances and gradients to objects. On the other hand, TSDFs are voxel grids representing the distance to the nearest surface along the ray direction from the center of the sensor and are commonly used for fast, flexible map representation. TSDF use a projective distance to produce surface meshes using zero-crossings, which means that TSDF representation is truncated at a certain distance reducing the computation of the process. These values can be positive if the voxel is inside the object, negative if it's outside the object and zero if it's on the object's surface [21].

The approach in Voxblox is to incrementally build ESDFs from TSDFs making it applicable for both mapping and surface reconstruction and also for planning while overcoming the shortcomings of occupancy maps. Occupancy maps compute the distance to obstacles at all points in a map but can't be dynamically changed and needed to know a priori the maximum size of the map. The proposed method in Voxblox is able to build ESDFs directly out of TSDFs and exploits the pre-existing distance information within the truncation radius. Additionally, the voxel hashing method allows the map to grow dynamically by allocating blocks of fixed size. On the other hand, the marching cubes algorithm generates a 3D surface mesh from volumetric data. This algorithm uses a collection of 3D data points and divides them into small cubes that contain values representing the data. These values define the shape of the surface that is represented [25].

Hydra leverages the algorithms developed in Voxblox and makes some improvements. The metric-semantic mesh is generated within a radius of eight meters, known as active window, that bounds the amount of memory. Additionally, Hydra assigns labels to the TSDF voxels representing zero-crossings, which indicates surfaces and are referred to as "parents". Furthermore, the ESDF holds the distance to the nearest obstacle and the parents that are closest to them. Then, the metric-semantic mesh is formed within this radius and as it moves out of this bound it is then passed to the Scene Graph Frontend [1].

Objects and Agents

The second layer constitutes a subset of objects and agents, where each object is described with a semantic label, a centroid and a bounding box, while agents are represented through a pose graph detailing its trajectory. Hydra performs object segmentation through individual Euclidean clustering of 3D metric-semantic mesh vertices, with each semantic class being grouped separately. Similar to Kimera, the Euclidean clustering is used to estimate the centroid and the bounding box for each potential object. When a potential object overlaps with an existing object node of the same semantic class they merge and they incorporate new mesh vertices into the previous object node. However, if the new object does not align with an existing object node, it is included as a new node [1].

Places

The third layer groups the places of a scene, in which each place represents the open spaces and the edges symbolize traversal paths. This layer firstly builds a skeleton diagram and then forms a sparse 3D graph of the environment by leveraging the work for micro-aerial vehicle planning [27]. Hydra incrementally builds a skeleton diagram that represents the topological structure using a Generalized Voronoi Diagram (GVD). GVD emerges as a result of the ESDF integration process and represent a collection of voxels equidistant to a minimum of two obstacles. This representation is gradually built through iterations that simplify the GVD into a subgraph of places that reduces the nodes and edges of the GVD. The result is a subgraph of places which represent the essential nodes that correspond to places and the edges that connect these locations. This graph preserves the key structural and spatial relationships among significant points of the environment within the active window. Then, it is passed to the Scene Graph Frontend, along with the metric-semantic mesh, as they exit the active window.

Rooms

The fourth layer builds upon the subgraph of places to define rooms. Each room is represented by a centroid and the edges establish links between contiguous rooms. Hydra innovates in the room segmentation process, which is guided by two main insights. Firstly, it utilizes dilation operation on the obstacles to gradually uncover rooms by closing apertures like doors. Secondly, this process will modify the sparse graph by erasing some places of the sparse graph. Therefore, Hydra performs the dilation operation in δ distances, progressively revealing distinct rooms and removing the nodes with distance smaller than δ as well as their corresponding edges.

Buildings

Lastly, the building node represents the entire building to which all the room are linked to. As mentioned in the different layers, edges establish relationships between entities of the same layer or even across different layers. Edges within the same layer represent the ability to move or traverse between different places or rooms. On the other hand, edges across layers can indicate a more complex relationship. For instance, an edge could show that mesh points are part of a specific object, that an object is located at a specific place or that a room is part of a larger building. Ultimately, edges are connections that model how different pieces of information or entities are related to each other, either within the same layer or across different layers of the system.

Architecture

This framework is designed with an architecture that combines low-level, mid-level and high-level perception processes to ensure an efficient operation in real-time, as shown in Figure 2.4.

This architecture aims to prevent slower and less frequent computations from hindering the execution of faster tasks. Hydra initiates its process by engaging in faster perception processes like 2D semantic segmentation and stereo-depth reconstruction that run at keyframe rate and tasks like feature detection and tracking that run at keyframe rate. Following the low-level section, the results are passed to the mid-level perception processes. These processes involve the gradual construction of different layers, the mesh and places layer and the object layer. Thus, the outputs of these modules are the latest mesh, places subgraph, objects and pose graph of the agent that the scene graph frontend receive as inputs. The frontend is the main output of this mid-level section and it continually updates the

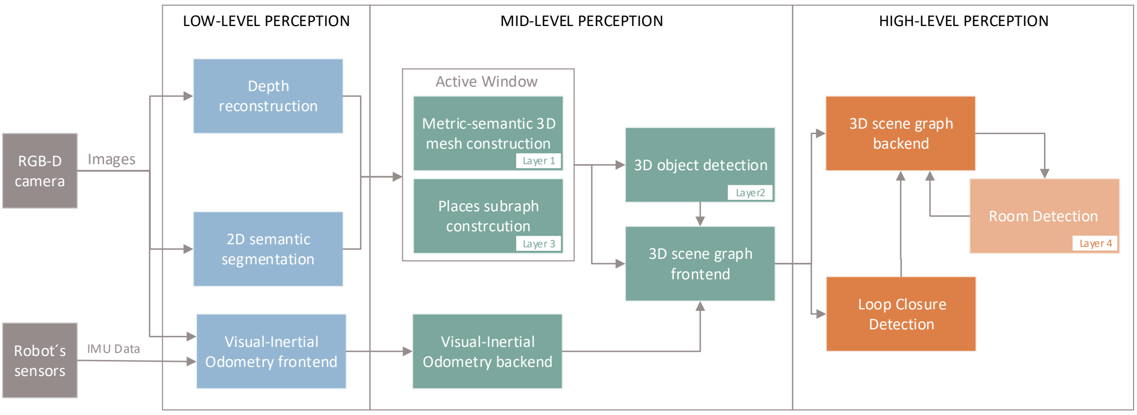


Figure 2.4: Hydra's architecture integrating low-level, mid-level and high-level processes for 3DSG construction [1].

3D scene graph. It creates an initial 3D scene graph that has not yet been optimized by combining the outputs from previous modules that progressively contribute to the scene graph upon exiting the active window. Lastly, the high-level perception processes like the loop closure detection, the scene graph backend optimization and the room detection run at slower rates. These tasks collectively construct a 3D scene graph that is coherent and persistent [1].

2.3 Smart buildings

The concept of smart buildings has evolved since the introduction of the term intelligent buildings in the 1980s. Initially, experts defined intelligent buildings as any building that could control its environment systems, such as heating, lighting and more, often managed by a computer system. This definition focused on achieving automation independently with minimal user interaction. Building on this concept, the term smart buildings intents to integrate and harmonize various aspects aiming to create a more holistic building system. In essence, smart buildings are comprehensive systems that combine intelligence, integration, control and innovative materials to promote adaptability, driven by energy efficiency, longevity and occupant satisfaction. These infrastructures aspire to create a more intelligent and efficient urban landscape within the broader context of smart cities. The future envisions smart cities as a response to growing urban population, demanding greater functionality from limited resources and stricter regulations [28].

2.3.1 Smart city taxonomies

The widely accepted concept of smart city revolves around the idea of introducing innovative technologies and applications across various aspects of city life. At the same time, smart cities aim to improve existing processes. To achieve this, they rely on the dynamic and ongoing process that work towards achieving additional improvements in urban areas. The term of smart city emphasizes the need to invest in different infrastructures for purposes like economic growth, improved quality of life, better resource management and participatory governance. In this context, smart cities are a comprehensive ecosystem that integrate a wide range of application domains and technologies, as represented in Figure 2.5.

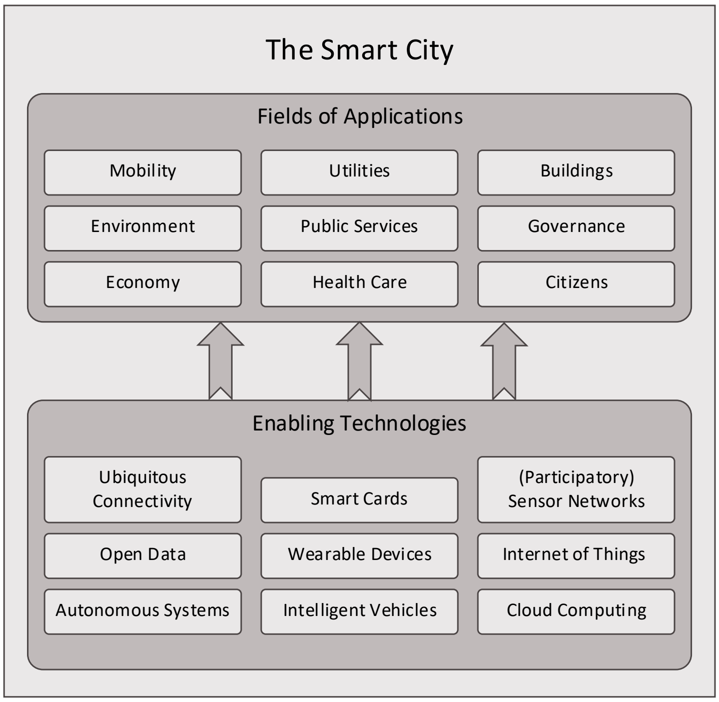


Figure 2.5: Smart city taxonomies [29].

Applications

Smart cities relies on the interconnection of different applications and technologies that aim to improve urban areas. The applications are classified into nine different areas: Mobility, Utilities, Buildings, Environment, Public Services, Governance, Economy, Health Care and Citizens, as shown at the top of Figure 2.5. These key applications work jointly in a collaboratively manner to enhance city efficiency and effectiveness. In this context, smart buildings focus on making residential and commercial constructions more energy-efficient and convenient by incorporating features such as structural health monitoring, lighting and heating regulation based on occupancy and the use of intelligent appliances for automation. Moreover, home automation systems enhance energy efficiency and comfort in buildings [29].

Enabling technologies

The technologies that enable these applications are classified into nine categories: Ubiquitous Connectivity, Smart Cards, (Participatory) Sensor Networks, Wearable Devices, the Internet of Things (IoT), Intelligent Vehicles, Autonomous Systems, Cloud Computing and Open Data, as shown at the bottom of the Figure 2.5. These enabling technologies frequently collaborate to drive a wide range of smart city applications and they provide the mechanisms, resources and support that applications need to operate. In this context, IoT involves connecting everyday objects, like household appliances and vehicles, to the Internet. These objects, equipped with sensors and actuators, can exchange data and enhance their functionality. Similarly, cloud computing involves outsourcing computational tasks to third party providers, offering scalability and efficient data analysis [29].

2.3.2 IoT Technology

IoT is a technology that enables the operation of diverse applications in the smart city domain. This technology denotes the technological advancements of the internet, enabling billions of devices to communicate and interact. It aims for Machine-to-Machine (M2M) communication, allowing devices to interact without human intervention. Smart devices play a pivotal role by constantly collecting and sharing data. The term ”smart” refers to their ability to make real-time decisions based on continuous data monitoring. These devices can range from sensors, actuators and cameras to household items and wearable devices, such as smartwatches, generating significant volume of data. The scope of smart devices extends beyond physical devices or objects to human behavior and information as well. This extensive network of interconnected smart devices requires for a robust infrastructure that overcomes the implementation challenges of this technology. Some of these challenges include network latency, where critical applications demand minimal network latency; the absence of a standardized platform and common architecture, as standardization is essential for growth and security; and scalability issues, as the system must expand with the increasing number of implemented application and devices while ensuring availability and efficient bandwidth management [30].

The volume of data that a system of smart devices is capable of generating and exchanging is immense. This large volume of data is also known as Big Data, which is an abstract concept that refers to the vast amount of data that needs to be processed rapidly. Big Data is necessary to gather, process and derive value from the data. Generally, the data collected in IoT scenarios lacks structure and requires a comprehensive analysis to derive meaningful insights. Big data complements IoT due to its ability to handle the substantial volumes of data from various sources. The integration of IoT and Big Data will impact the way data is stored, processed and analyzed. Additionally, cloud platforms can assist in the offloading of large amounts of data. They offer resources for processing and storing the massive volume of data produced by IoT devices. These platforms provide the necessary infrastructure and computing capabilities to handle the influx of data efficiently. Moreover, the real-time requirement in IoT scenarios make cloud computing less suitable. Fog/edge computing brings computation closer to the data sources, reducing processing time and enabling quicker responses [30].

IoT Architecture

The implementation of a standardized network infrastructure in the IoT scenario aims to integrate different networks. The necessity for this integration arises due to the fact that multiple heterogeneous networks must coexist and interact with each other. Within this infrastructure, all IoT applications would be capable of sharing network resources and information to maximize the potential of every interconnected element within the network. As a result, all applications would have the ability to easily communicate and efficiently share resources. Furthermore, the implementation of a generalized network infrastructure holds the potential to reduce the cost of network deployment.

In the research literature, numerous IoT architectures have been proposed, ranging from the fundamental three-layer model to more intricate multi-layered designs. The foundational architecture employed in IoT is structured into three layers, including the perception layer, the network layer and the application layers as shown in Figure 2.6 [31].

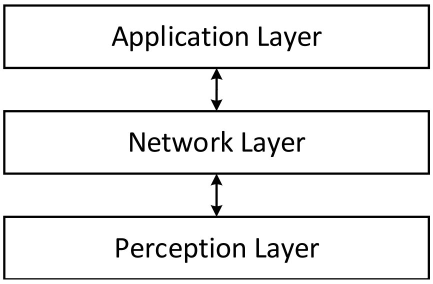


Figure 2.6: IoT basic three layer architecture [31].

Positioned at the bottom layer of the IoT architecture, the perception layer is responsible for the interactions with the physical world. It is also known as the sensor layer and it manages the measurements, collection and processing of the state information of the physical devices through the smart devices. Positioned within the middle layer, the network layer is responsible for the transmission and integration of the data. This layer is also known as the transmission layer and it receives the information from the perception layer to redirect it. It selects the appropriate route for the information to reach the intended destination, ensuring seamless communication between different components of the IoT system. Moreover, this layer integrates diverse devices, communication technologies and performs data transmission among various entities using different technologies and protocols within networks. Positioned as the top layer, the application layer is in control of the data received from the network layer and uses it to deliver the necessary services or functionalities. This layer is also known as the business layer and can provide services such as storage or analytical tasks. Each layer communicates through gateways, which are middleware that provide a bridge between layers that may use different communication protocols and ensures bidirectional communication, allowing the flow of information between layers.

2.3.3 Cloud and fog/edge computing

Fog/edge computing is a distributed architecture that offers computing services between central servers and end-users, enabling real-time processing of substantial IoT generated data at the network edge. The integration of fog/edge computing with IoT is envisioned as the forthcoming infrastructure for IoT development. This integration improves the bandwidth and energy consumption issues related with cloud computing, while delivering faster responses.

The term cloud computing refers to the process of computing in the cloud, while edge computing refers to computing at the network's edges. The main distinction between these two concepts lies in where data analysis takes place. In edge computing, data storage, analysis and processing occur in real-time near the device collecting the data. This reduces latency, conserves bandwidth and facilitates rapid and effective processing of data. A limitation of this approach is that it processes only locally collected data, making it challenging to apply extensive Big Data analytics. On the other hand, cloud computing allows the utilization of various services, such as storage and servers, through internet connectivity. The cloud enables gathering massive amounts of data from various sources and analyzing them in diverse ways to make intelligent decisions. In this scenario, data travels from the source device to the cloud, where it undergoes processing and analysis. Processed data, if needed, is then returned to the devices. This data transmission process introduces a significant challenge known as network latency, which causes delays in the communication. Although this delay might be minimal, it can be critical for certain applications. Cloud computing compensates this drawback with power and capacity, relying on a scalable data centre infrastructure to expand storage and processing capabilities as required.

These two technologies are not mutually exclusive, as each brings benefits to specific applications and can effectively complement one another. Edge computing is optimal for applications where slight delays are critical, processing and analyzing data closer to the source device. Conversely, cloud computing enables large-scale data analysis that is not feasible at the network edge. By incorporating both computing approaches into a single network infrastructure, it is possible to maximize the potential of both approaches while minimizing their limitations. A representation of this configuration sis shown in Figure 2.7.

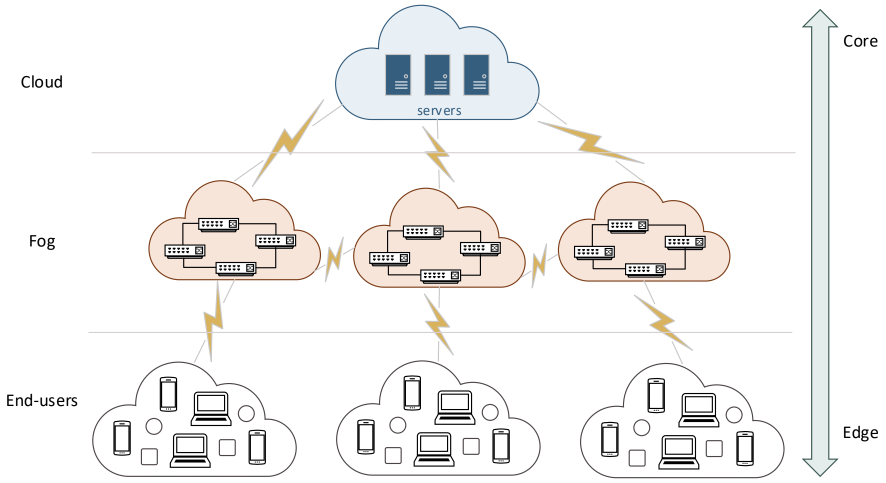


Figure 2.7: Cloud and fog/edge computing [30].

Fog/edge nodes can be any device that is equipped with storage, computing capabilities and network connectivity and they can be deployed at any location to gather data for different IoT applications. The structure shown in figure 2.7 shows how devices are interconnected. Based on the priority of the data, the different types of IoT data can either be analyzed directly at the edge or transmitted to the cloud for processing. Data that requires quick responses have a higher priority and is processed closer to the IoT devices. Alternatively, less time-sensitive data has less priority and can be rerouted to aggregation nodes for processing. However, there are some challenges to integrate fog/edge computing with IoT due to the limited computing and storage capability of each fog/edge node. These challenges include the efficient management of fog/edge computing resources between end devices and fog/edge nodes and among fog/edge nodes. Firstly, each end user is defined with a satisfaction function to optimally allocate the resources between end devices and fog/edge nodes. This function aims to maximize the overall satisfaction of users and is calculated based on the available resources that the nodes have, the resources that are allocated for each user and their priority. Secondly, managing resources among end nodes is mitigated by sharing of computational resources among the neighbor nodes. The interconnection of all end-nodes and fog/edge nodes facilitates the utilization of available resources from other nodes that have available capacity [30].

3. Privacy Threat Modelling

In this chapter, I undertake a comprehensive exploration of threat modeling with a specific focus on privacy concerns. The primary goals of this project are to understand and address potential threats to privacy within the context of 3DSG and to propose effective Privacy-Enhancing Technologies (PETs). To achieve these objectives, the LINDDUN methodology is selected due to its focus on privacy concerns. Within this chapter, detailed overviews of the privacy threats and the methodology of this approach are provided.

3.1 Introduction to threat modeling

Threat modelling is the process of systematically identifying and assessing potential threats that could negatively impact a particular system, application or network. Any undesirable event that has the potential to compromise the security or integrity of a system is considered a threat. Threat modeling focuses on understanding the nature of the threats and the impact they may have while considering potential vulnerabilities. The aim is to ensure data privacy by proposing specific PETs to address threats and protect the confidentiality of sensitive information. This process can be applied in various scenarios, including systems, networks, or business processes and it enables informed decision-making continuously through the development process. Applying threat modeling during the initial stages of the development process can help achieve cost savings. However, it is also beneficial for practitioners to implement threat modeling continuously as they add more complexity and acquire a greater understanding of the system [32][33].

3.1.1 Threat modeling frameworks

Threat modeling is a versatile practice applicable to a wide range of systems, each potentially exposed to unique risks and threats. Typically, threat modeling methods involve the creation of a system abstraction, profiling potential attackers considering their goals and methods and a catalog of possible threats. There exist a variety of approaches to apply threat modeling, each tailored to address specific aspects of security or privacy. The different methodologies range from technical, diving into the intricate details of a system's architecture, to more people-centric, considering the human actions or behaviors that can influence security vulnerabilities. Some of these methodologies are listed and briefly explained below [34]:

STRIDE

STRIDE is a security-oriented threat modeling framework that focuses on identifying and grouping threats based on six categories: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service and Elevation of Privilege. This framework follows an iterative process that classifies system assets into the security categories to formulate distinct threat scenarios and suggest appropriate mitigation strategies. This mature approach is straightforward to apply but can be very time consuming as the quantity of potential threats escalates rapidly with the growth of complex systems [34][35].

PASTA

The Process for Attack Simulation and Threat Analysis (PASTA) is a risk-focused threat methodology that is divided into seven stages. This framework tries to connect the business objectives with the technical requirements of a system, while promoting the cooperation of stakeholders. This methodology adopts an attacker's perspective to identify and analyze the threats, assigning ratings to prioritize the mitigation of the most critical threats early in the process [34].

LINDDUN

LINDDUN is a privacy focused threat modeling methodology that stands for seven privacy threat categories: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of Information and Unawareness. Similarly to STRIDE, this framework systematically analyzes the system assets from the viewpoint of the threat categories to formulate threat scenarios and finally propose suitable PETs. This systematic approach is mainly used in the context of web applications and comprises six steps, which fall into two categories, the problem space and the solution space [34].

Trike

Trike is a comprehensive threat modeling methodology that emphasizes security concerns from a risk management perspective. This framework systematically assesses and mitigates security threats within a system by first defining its components and creating interaction matrices. In this approach, threats are categorized as either elevation of privilege or denial of service and it examines CRUD actions (Create, Read, Update, Delete) to assess the risk of potential attacks [34].

VAST

The Visual, Agile and Simple Threat (VAST) is threat modeling approach that focuses on cybersecurity for large organization systems. This method relies on automation, integration and collaboration to offer scalability for organizations. VAST recognizes two types of models, the application threat models and the operational threat models. The former focuses on creating a representation of the system and the latter consider the attackers point of view based on the representation [34][36].

OCTAVE

The Operational Critical Threat, Asset and Vulnerability Evaluation (OCTAVE) is a risk-based assessment methodology that focuses on assessing organizational risks through system reviews. It is primarily focused on operational risks while excluding technological risks. The three phases of this approach involve creating threat profiles, evaluating information infrastructure vulnerabilities and formulating security strategies and plans to protect critical assets [34].

The choice of which method to employ depends on the unique characteristics and objectives of the system under analysis. There is not a framework that can perfectly fit to a particular use case. It is essential to select one that aligns most closely with the needs and goals. Moreover, different methods can be combined to provide a comprehensive assessment of potential threats. In this particular study, the LIDDUN framework stands out as the unique approach mainly centered on addressing privacy rather than security concerns.

3.2 LINDDUN Privacy Threat Modelling

As introduced before, LINDDUN is a methodology for privacy threat modeling that supports systematic elicitation and mitigation in software systems. This model-based approach focuses on privacy concerns and it is an acronym for the privacy threat categories it supports: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness and Non-compliance, which are described in 3.2.1. This methodology is designed to provide structured guidance for the privacy threat modeling process through a step-by-step method described in 3.2.2.

LINDDUN provides a systematic and practical approach for evaluating a software system's privacy posture, identifying privacy flaws and recommending PETs through the use of its proposed threat types, mapping tables, threat trees, mitigation strategies and privacy-enhancing solutions. Thus, this privacy focused framework guarantees an extensive privacy knowledge support, which ensures that the entire privacy assessment process is thorough, covering all aspects and well-documented. Additionally, this methodology is compatible with STRIDE since both STRIDE and LINDDUN share similar principles. Both frameworks start from the system model, allowing for concurrent security and privacy assessment. This alignment simplifies the integration of the LINDDUN framework into the risk assessment phase of the design process for practitioners. It's specific emphasis on privacy issues makes it an effective choice for understanding and mitigating potential privacy threats within the context of 3DSG. It also ensures a smooth incorporation into existing security assessment practices [7][37][38].

3.2.1 Threat categories

As mentioned before, LINDDUN is an acronym for the privacy threat categories it supports. The seven threat categories are described in detail below [37][39].

Linkability

Linkability refers to the process of connecting different data elements in a manner that can reveal additional information about an unknown individual or group. This threat involves matching distinct items of interest (IOI) that are related to the same user, which can lead to identifiability, which is described later on. In addition, linking can be extended to data belonging to multiple individuals by identifying common attributes or properties. This can result in revealing information about the group as a whole and can lead to group profiling and inference. Inference is a technique applied to linkable data that can generalize relations based on properties and can be used to social discrimination.

The concept of linkability derives from the definition of unlinkability [40], which denotes the inability from the attacker's perspective to relate two IOI. Therefore, linkability is the negation on unlinkability and can be described as the ability to differentiate whether two IOIs are related or not, without requiring knowledge of the subject associated with the linked IOI.

Identifiability

Identifiability refers to the capability of identifying a subject within a larger group of subjects that share similar attributes or properties through an IOI. The concept of identifiability is closely related to anonymity and pseudonymity [40]. In the context of an attacker that attempts to identify a subject, anonymity refers to the inability of the attacker to distinguish the subject within a set of subjects called the anonymity set. Beyond that, in order to capture the possibility to quantify anonymity, this term refers to the extent to which the attacker cannot sufficiently identify the subject within the anonymity set. The purpose of anonymity is to hide the link between the identity of a subject and a piece of information. In the same context, pseudonymity refers to the use of pseudonyms as identifiers of a subject other than the subject real names.

Data can be considered de-identifiers when it does not contain identifiable information, such as social security number or birth date. However, the use of pseudo-identifiers such as birth year instead of birthdate, or city instead of full address, can potentially result in identification. As mentioned previously, linkability poses a risk of identifiability by associating multiple pseudo-identifiers together which are not able to individually identify a subject. However, by linking them together, the anonymity set decreases, which increases the likelihood of identifiability.

Non-repudiation

Non-repudiation involves having irrefutable evidence regarding a certain action or fact that whether confirms or denies its occurrence and attributing it to a specific individual. This threat holds an individual not able to deny their involvement in a specific claim as a consequence of the data that has been collected or shared or actions taken by the individual. Non-repudiation is closely related to plausible deniability, a term used when a subject can deny any involvement in an action or decision [37]. Therefore, non-repudiation leads to the loss of plausible deniability, being these two terms mutually exclusive. Additionally, non-repudiation threats will be increased with identifying and linking threats, since the more information that is associated with a individual, the more difficult it becomes to deny their involvement.

Detectability

Detectability refers to the ability to determine the involvement of an individual through observation of the exitance of relevant information, without its disclosure. Detecting is built upon the observed communication, observed application side-effects, or evoked system responses that may leak the existence of certain elements. Through observation, the attacker deduces additional information and this threat does not require direct access to the data.

Detectability is closely related to undetectability and unobservability [40]. From an attacker's perspective, undetectability refers to the inability to accurately determine the existence of an IOI. On the other hand, unobservability pertains to the undetectability and anonymity of the subjects involved in the IOI, even among each other. Additionally, detectability is related to linkability and identifiability since the deduced information about an individual can extract more information.

Disclosure of information

Disclosure of information refers to exposing sensitive information to someone not authorized to see it. Known or unknown third parties may use the information for unauthorized purposes. This threat can arise from collecting, storing, processing, or sharing excessive personal data. Data disclosure threats represent cases where attackers disclose personal data either explicitly or implicitly. Attackers intentionally and by design engage in explicit disclosure, while they indirectly and unintentionally cause implicit disclosure. This threat closely relates to confidentiality, which involves hiding information or managing data content release. Although confidentiality is a security property, it is also important for preserving privacy properties, such as anonymity and unlinkability.

Disclosure threats center around four primary characteristics: unnecessary data types that increase data susceptibility to misuse; excessive data volume, gathering higher amounts of data can increase the risk; unnecessary processing, which involves further data treatment that can increase the risk; and excessive exposure, that widens access to information by exposing it to more parties.

Unawareness

Unawareness occurs when organizations inadequately inform individuals about the handling of their personal data. This threat includes a lack of transparency, where organizations fail to properly inform the data subjects about the collection and/or processing of their personal data or that of others; a lack of feedback, where data subjects remain unaware of the potential privacy implications of sharing their personal information; and a lack of control, where data subjects cannot access or control their data.

Non-compliance

Non-compliance refers to deviating from legislation, regulation, standards and/or policy, which leads to insufficient risk management. This threat assessment determines whether the system's policies and the user's consent are being properly implemented and enforced. It is important to consider non-compliance in the broader context of risk management, including legal and cybersecurity risks and to evaluate its relation to other privacy threats, such as linking, identifying, non-repudiation, data disclosure and unawareness. In LINDDUN, non-compliance threats mainly focus on threats that derive from other privacy threat categories.

In summary, these categories can be described as follows.

Table 3.1: LINDDUN threat categories [39].

Threat categories	Description
Linkability	Associating data items or user actions to learn more about an individual or group of individuals.
Identifiability	Learning the identity of an individual.
Non-repudiation	Being able to attribute a claim to an individual.
Detectability	Deducing the involvement of an individual through observation.
Data disclosure	Excessively collecting, storing, processing, or sharing personal data.
Unawareness	Insufficiently informing, involving, or empowering individuals in the processing of personal data.
Non-compliance	Deviating from security and data management best practices, standards and legislation.

3.2.2 Methodology step-by-step

LINDDUN provides a systematic approach to privacy assessment and can be divided into three main steps. Firstly, it produces a system model by decomposing the processes, entities, data stores and data flows to gain a deeper understanding of the systems' working. This is achieved by designing a Data Flow Diagram (DFD) that serves as a graphic representation of the system under analysis. Once the system is modeled, the next step involves systematically iterating over each DFD element to elicit potential privacy threats. Finally, suitable solutions are proposed to manage and mitigate the identified threats. The solution must address each specific threat in a manner that minimizes its impact while ensuring the continued protection of the data subject's privacy. In greater detail, LINDDUN involve six steps which are categorized into two categories as shown in Figure 3.1.

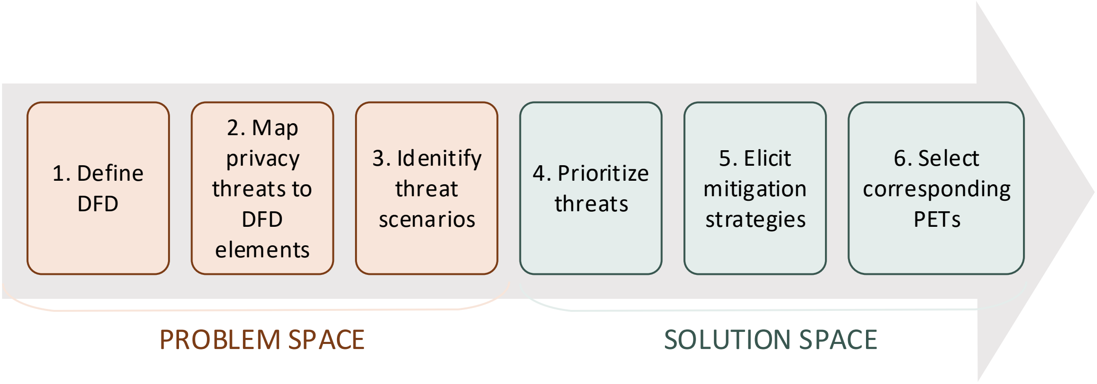

1. Define DFD
2. Map privacy threats to DFD elements
3. Idenitify threat scenarios
4. Prioritize threats
5. Elicit mitigation strategies
6. Select corresponding PETs
PROBLEM SPACE
SOLUTION SPACE

Figure 3.1: LINDDUN methodology steps[37].

The first category is centered on the problem space and aids in identifying privacy threats. The second category focuses on the solution space and aims to address the threats identified in the previous category while providing privacy solutions for them [37].

Define data flow diagram

The first step is to create a DFD that models the system under observation. A DFD is a conceptual, structured and graphical representation that decomposes an Information System into its main components. This representation sets the basis for the future analysis and any inaccuracies in the diagram could lead to inconsistent results in the analysis of the actual system. The components of this representation can be classified into four major building blocks: Entities (E), which can either be a source or a destination of data; Data Stores (DS), which is a logical repository of data; Data Flows (DF), which represent the movement of data across the system; and Processes (P), which represent a unit that operates on the data. These blocks aim to represent how the data moves within the system, the user-interaction points and the trust boundary, which describes the limits of the system. The trust boundary separates the internal system from the external parties and can be interpreted as the point of interaction between parties with different privileges [41][42].

DFDs can represent systems at different levels of abstraction, which can go from a high-level representation to a more in detail representation by partitioning the system functions. The level of detail of the elicited threats will be influenced by the granularity of the DFD. The distinct levels for the DFDs are numbered as level-0, level-1 and level-2 DFD. At the requirements level, developers create the level 0 DFD, which corresponds to a context diagram that represents the system as a single main process. This process is connected to the external entities such as users of the system and third parties. In the subsequent levels, the main process is gradually decomposed into multiple processes that represent the main function of the system. The level 1 DFD aims to give a more detailed representation of the internal processes, while the level 2 DFD dives deeper into the subprocesses identified in the level 1 DFD. This level of detail can be helpful in planning or documenting the system's specific operational aspects [43].

Map privacy threats to DFD elements

The second step consists of mapping privacy threats to the DFD elements. As presented in Table 3.1 the LINDDUN methodology considers seven types of threats, which helps categorize the threats. The outcome of this step is a table that associates every DFD element to each privacy threat category. Every 'X' in this table displays the susceptibility of the DFD element to the corresponding threat category. In general terms, Table 3.2 shows a template that specifies the threat categories that are relevant to each element type in a DFD.

Table 3.2: Mapping template for privacy threats to DFD element types [37].

Threat categories	**Entity**	**Data Flow**	**Data Store**	**Process**
Linkability	X	X	X	X
Identifiability	X	X	X	X
Non-repudiation		X	X	X
Detectability		X	X	X
Data disclosure		X	X	X
Unawareness	X			
Non-compliance		X	X	X

When implementing this template, practitioners need to tailor it to the specific system being analyzed. This template acts as a guide, ensuring the consideration of all privacy threat categories for each DFD element. As a result, practitioners must document every 'X' individually, considering it as a potential threat directed to a specific DFD element. However, experts can combine multiple 'X' that that are related to the same type of DFD element and that pose comparable repercussions with equivalent priority level as a unique threat. As en exception, non-compliance threats can be considered to the entire system since they are more generic, allowing practitioners to handle them together.

Identify threat scenarios

The third step is eliciting privacy threats, which can be considered the core execution step of the LINDDUN methodology and that will result in the documentation of the misuse cases. This step starts with refining threats via threat tree patterns inspired by the Secure Development Lifecycle (SDL) [44]. LINDDUN threat trees break down each threat category into more specific characteristics for a more detailed and comprehensive view of the threats. Overall, they are a structured guide to help designers in considering privacy conditions within the system and are regularly updated. They provide a corresponding threat tree that illustrates the specific vulnerabilities that can be targeted and exploited to carry out a privacy attack scenario and that must be considered for every 'X' marked in the mapping template.

Additionally, practitioners should document when they trust an element of a system to behave as expected in accordance with assumptions. These assumptions are decisions that help evaluating the relevance of a threat or category within the system during the elicitation phase. It is essential to keep documentation and provide a link to the corresponding misuse case for traceability. Moreover, assumptions can serve as domain restrictions, constraining the scope of the LINDDUN analysis and potentially reducing the number of threats to examine through the use of broader assumptions. Finally, the outcome of this step is a collection of threat scenarios documented as misuse cases. A misuse case can be understood as a use case from the attacker point of view. For the documentation, LINDDUN provides a threat description template that I have adapted for this work as presented in Table 3.3.

Table 3.3: Description of threat scenarios template.

Summary	*Brief description of the scenario.*
DFD elements involved	*Relevant components of the DFD.*
LINDDUN properties	*Categories that apply to the threat.*
Assets involved and consequences	*Brief overview of the assets involved in the threats and the consequences related to them.*
Priority	*Indicate how important you consider the threat, based on the impact and likelihood of the threat, as described in 3.2.2.*

Prioritize threats

The LINDDUN methodology may generate many documented threats that need to be addressed. To move forward and find appropriate mitigation strategies, it is necessary to prioritize the identified threats. LINDDUN does not offer direct assistance with conducting risk analysis. Instead, it refers to established techniques such as OWASP's Risk Rating Methodology, described in 3.3 [45], Microsoft's DREAD [46], NIST's Special Publication 800-30 [47], or SEI's OCTAVE [48]. These techniques use the information from the misuse cases, including the assets involved for impact and the attacker profile and basic/alternative flows for likelihood.

Alternatively, practitioners prioritize threats by selectively focusing on specific threats during the identification phase and only the most critical ones are usually considered for inclusion in the solution design. Typically, you can calculate risk as a product of the likelihood of the attack scenario and its impact, as shown in Equation 3.1. Then, experts use this product to rank the risks according to their priority, with higher priority indicating higher risks.

$$Risk = likelihood \times impact \tag{3.1}$$

Focusing on the documented threats with higher likelihood and impact can save time and effort, while not considering threats that are more unlikely.

Elicit mitigation strategies

After identifying and prioritizing the threats, LINDDUN proceeds to find suitable solutions to prevent or resolve these threats. This methodology aids this process by providing corresponding mitigation strategies, which will then be connected into a technical solution, enabling the transformation of privacy threats into effective and tailored PETs.

The mitigation strategies helps thinking about the general approach, instead of the specific solution for each threat and can be seen as an intermediate step to finding appropriate PETs. They offer a structured categorization of commonly used methods to address privacy threats, providing a systematic approach for identifying and resolving those. The mitigation taxonomy focuses on two major strategies, as shown in Figure 3.2.

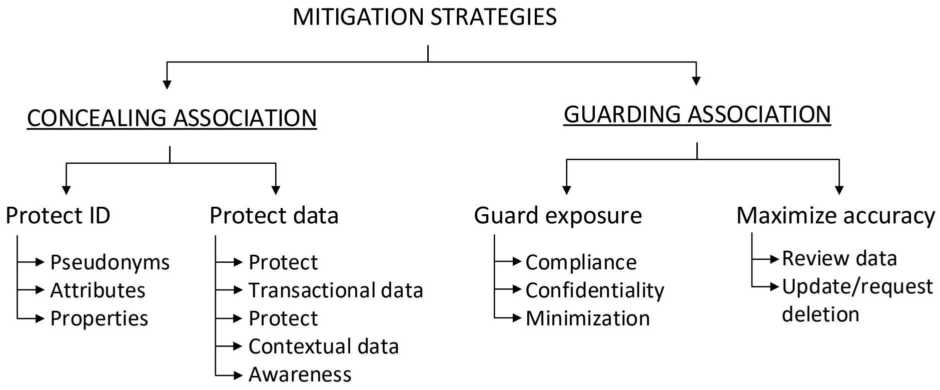


Figure 3.2: LINDDUN mitigation strategies [37].

The first strategy is a proactive approach that involves concealing the associations between users and their transactions and personal information. This strategy focus on ensuring that only necessary information is shared with the system. The second strategy is a reactive approach, which aims to limit the damage by controlling the associations after disclosure and minimizing their exposure. As a result, it is necessary to determine if applying the solution can be made proactively before collecting the data, or reactively at storage time. Once that it done, it is necessary to find a suitable strategy at that stage.

The selection of mitigation strategies can vary depending on the type of DFD element. For instance, threats related to entities and data flows correspond to concealment of data, while threats related to data stores correspond to guarding the exposure of association. Mitigation strategies related to entities aim to protect the identity and conceal the data and data flow threats can be resolved by protecting data before and during the communication. On the other hand, threats related to data stores, that is data that has already been collected and stored, can be mitigated by ensuring confidentiality or minimizing the data. Additionally, a data subject has the ability to modify the data that has been collected about them, which involved the non-repudiation threat category. Furthermore, both branches of the taxonomy can be associated with unawareness threats. Users need to be aware of the potential implications of sharing information and must remain informed throughout the entire process of the data life cycle.

Select corresponding PETs

The last step involves identifying appropriate PETs to provide a more precise selection of solutions. Building upon the mitigation strategies from the previous step, LINDDUN then narrows down the range of possible solutions. This methodology provides an initial table categorizing PETs according to the mitigation strategies [37]. These categorizations aid in the process of identifying the most suitable PET to effectively address an individual threat, considering the complexity of this task due to the amount of possible PETs.

3.3 Risk assessment

Risk assessment is a structured process to identify potential threats and analyze their causes and consequences. The aim of this process is to describe and prioritize the risks, while evaluating their likelihood of occurrence and their severity. In the assessment, multiple scenarios with various failures are considered, providing a comprehensive understanding and knowledge of the system's failure modes. Primarily, there are two main categories of risk assessment, quantitative and qualitative [49].

3.3.1 Qualitative assessment

Qualitative risk assessment involves evaluating assets for vulnerabilities and assessing the probability of a threat using non-numerical values. It uses a relative scale with values like Low, Moderate and High to measure the impact and likelihood of a threat. This approach does not require complex calculations, but may lead to ambiguous classifications, making later reassessment challenging. It simplifies risk evaluation but lacks detailed information for precise categorization.

3.3.2 Quantitative assessment

Quantitative risk assessment is a thorough method that involves numerical assessment of likelihood and impact, typically measured in frequency, like incidents per year. This approach aims to quantify the damage by assigning economic cost to vulnerabilities and threat events. It identifies the organizations risk based on the financial impact and it is commonly used in financial institutions and insurance companies. This method requires a more complex, expensive and more time to achieve.

3.3.3 OWASP Risk Rating Methodology

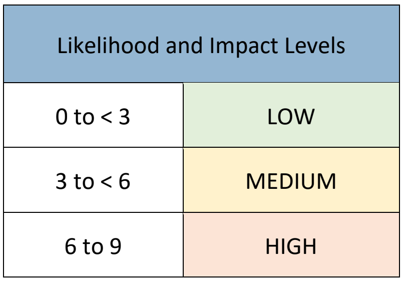
The Open Web Application Security Project (OWASP) is a multinational non-profit organization that aims to improve software security and promote awareness in the field. The OWASP Risk Rating Methodology (ORRM) helps to evaluate the severity of the risks and make informed decisions on how to manage them by adopting a six steps approach. This framework can be adapted for each organization needs, while balancing simplicity and sufficient detail. However, it is not necessary to be overly precise in this estimate, instead a general classification of low, medium, or high is sufficient [50].

The first step is to identify a security risk and rate it by gathering information about a threat agent, the attack, the vulnerability and the impact considering the worst-case that will result in the highest overall risk.

The second step is to evaluate the likelihood of the identified risk considering two sets of factors. The first set of factors is related to the attacker involved from a group of possible attackers. These factors include skill level, motive, opportunity and size. The second set of factors is related to the vulnerability involved. These factors include ease of discovery, ease of exploit, awareness and intrusion detection and they assume the attacker from the previous step. Each factor has a likelihood rating from 0 to 9, which are used to estimate the overall likelihood later.

The third step is to estimate the impact of a successful attack, which can be categorized into two types, technical impact and business impact. Technical impact factors are used to estimate the magnitude of the impact on a system. The factors to consider for technical impact include the loss of confidentiality, loss of integrity, loss of availability and loss of accountability. On the other hand, business impact factors are used to determine the potential impact on the business and they require a deep understanding of what's important to the company running the application. Common areas to consider include financial damage, reputation damage, non-compliance and privacy violation. However, business impact is more significant, but it may be challenging to obtain all the information needed to assess the business consequences. Therefore, providing detailed information about the technical risk can aid the appropriate business representative in making an informed decision. Each factor in impact assessment has a list of options with a corresponding impact rating from 0 to 9, which will be needed later to estimate the overall impact.

The fourth step is to determine the severity of the risk by combining the likelihood and the impact levels from the previous steps. The severity of the risk is then categorized as low, medium, or high, dividing the scale from 0 to 9 into three parts as shown in Figure 3.3.



Likelihood and Impact Levels	
0 to < 3	LOW
3 to < 6	MEDIUM
6 to 9	HIGH

Figure 3.3: Scale for the severity of the threats. [50]

The fifth step involves deciding which risks to prioritize for fixing, with the most severe risks taking priority. The outcome of this step will be a prioritized list. Lastly, the sixth step is to customize the risk rating model to ensure its effectiveness. There are several ways to customize the model such as adding factors, customizing options and weighting factors. This step helps to better represent what is important to a specific business and to emphasize the factors that are more significant. By customizing the risk rating model, businesses can create a framework that produces results that match people's perceptions about what is a serious risk, thereby promoting its adoption.

4. Problem Space

The following two chapters present the results obtained from applying the LINDDUN privacy threat modeling to Hydra's pipeline in the context of a smart building. For this purpose, the steps presented in Section 3.2.2 are followed.

This chapter is organized in three sections. The first section provides a comprehensive overview of the different components and different data types that have been considered for the modeling of the system. The second and third sections follow the problem-oriented steps to identify the possible privacy threats in the system for the level-0 and the level-1 DFD, respectively. The initial three steps of the LINDDUN methodology are focused on the problem space. Firstly, the system is modeled using a DFD. Following that, a threat mapping is carried out based on the model. Lastly, several threat scenarios are identified for both the level-0 and level-1 DFD.

4.1 Modelling

Before initiating the LINDDUN process, it is necessary to identify and specify the components and types of data that will be used in the smart building context. This information will play a vital role in defining the DFD, which will add the data flows and trust boundaries within the system. The successful integration of Hydra into a smart building context relies on the combination of diverse hardware and software components, which are necessary dependencies to enable real-time experiences [51][52].

- IoT sensors and actuators: IoT sensors are devices that detect and collect information from the surrounding environment within a specific range. IoT actuators are devices that produce some type of stimulus or action based on the received information from the sensors. Together, sensors and actuators form a network of connected devices that can interact with the physical world and each other, providing valuable data and automation capabilities. Hereafter, the IoT sensors and actuators will be referred to as IoT devices.
- RGB-D camera and robot's sensor: Hydra's process require real-time information from the environment, which is gathered by an RGB-D camera capturing images of the surroundings along with an integrated sensor in a robot or a wearable device, which provides IMU data that describes the orientation, velocity, and position of the robot within its surroundings.
- Edge and cloud servers: As the amount of data collected from IoT applications grows and spans over large geographical areas, efficient storage, processing, and analysis become increasingly important. While cloud computing provides computing and data storage services over the internet, fog/edge computing can extend it to be closer to the devices it supports. Fog/edge computing can provide computing and storage services to devices at the edge of the network to assist in processing tasks with higher computational requirements, instead of doing all the computation in the IoT devices.

- User devices: Users or building operators can interact with the building system and control certain features of their environment through their personal devices, as well as viewing information such as the operational status, current energy consumption, current emission, historical trends, etc. Supplied with this knowledge, users can make informed decisions that enable the building system to adjust the conditions accordingly while optimizing efficiency and with a focus on reducing energy consumption for sustainability due to rising energy costs.

Additionally, other stakeholders present in the context of a smart building are the service providers. This external party provides the smart building with different services that ensure the proper operation of the buildings systems and resolving any issues that may arise. All these components are interconnected and are constantly exchanging information. Given the vast amount of information being collected, transmitted, aggregated, and processed, both intentionally and unintentionally, it is crucial to clearly define the types of data used in the system and which vendors are utilizing it [51][53].

- Sensor data: It refers to the information collected and transmitted from the IoT devices. This data contains information of the surroundings and the metadata that characterizes a particular device.
  - Environmental data: Data related to the building's physical surroundings such as temperature, humidity, air quality, lighting, and energy consumption. This type of data is essential for optimizing building performance, enhancing occupant comfort and productivity, and minimizing energy usage for saving costs. Additionally, this information includes timestamps of when it was recollected.
  - Sensor metadata: Metadata regarding an IoT device can be differentiated between the observed/dynamic metadata, which describes the behavior of the sensor data, and device/static metadata, which describes the device and its parameters.
- User data: It refers to the PII collected and processed by the IoT sensors. PII is any data that could potentially identify a specific individual by containing sensitive information of the subject or bystanders.
  - Location data: Data related to the location of the data subject or bystanders and their environment.
  - Biometric data: Data related to measurable physical or behavioral characteristics of the data subject or bystanders in their environment. Examples of biometric data include facial features, iris or retinal patterns, voiceprints, and gait [54].
  - Identity data: Data containing unique references to the identity of the data subject or data that include revealing attributes which support the identification of the data subject.
- Service provider data: It refers to information from external sources that are not related to the building itself but can enrich the data available for smart building operations. For example, by incorporating weather forecast data into the building's control system, the system can anticipate changes in temperature or humidity and adjust heating and cooling accordingly, reducing energy consumption and costs.

In addition, sensitive information related to a smart building, that can be either a household or a workspace, should also be considered. This information may include sensitive details such as floor plans, occupancy patterns or personal information of individuals associated with the building. This data should be handled carefully to preserve the privacy and security of individuals and assets involved.

4.2 Level 0 DFD

Firstly, I present the level-0 DFD design. This representation illustrates the integration of Hydra as a subprocess within the context of a smart building. It corresponds to a context diagram that aims to analyze potential threats resulting from this novel technology's incorporation in this environment.

The goal of this section is to gain a comprehensive understanding of the privacy challenges posed by Hydra's integration into smart buildings, while also considering the inherent threats that smart buildings present. Examining the privacy implications arising from novel technologies and their potential usage in real scenarios serve as an early risk mitigation effort and provides guidance for responsible and secure advancements.

4.2.1 Define Data Flow Diagram

Initially, a level-0 DFD is presented in Figure 4.1, using the DeMarco notation. This representation relates all the components involved in a smart building operation, as well as the integration of Hydra. These interactions will be included in the threat model, but it will not include the architectural and logical design of these dependencies.

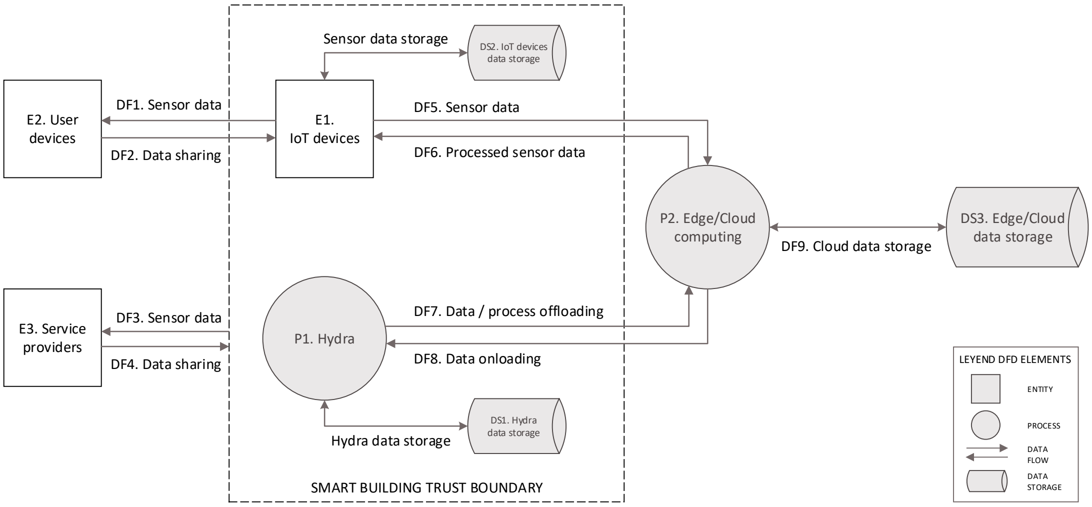


Figure 4.1: Level 0 Data Flow Diagram.

In the context of a smart building, the primary entities are the IoT devices, including both sensors and actuators, which engage in various forms of communications with external entities beyond the trust boundary. The external entities may include user devices, service providers and edge/cloud domain, and they are involved in various types of communication using several types of data. Additionally, Hydra's process is also integrated as a part of the smart building domain. The exchange of information is illustrated in Figure 4.1 and explained in Table 4.1.

Table 4.1: Description of data flows between entities in the level-0 DFD.

Source	**Data Flow Description**	**Destination**
E1. IoT devices	DF1. Sensor data containing environmental data from the user surroundings is sent as requested by the user.	E2. User devices
E2. User devices	DF2. User devices share data in the form of requests and information to modify environmental conditions.	E1. IoT devices
E1. IoT devices	DF3. Sensor data containing environmental data from the user surroundings is sent periodically to the external providers.	E3. Service providers
E3. Service providers	DF4. The smart building domain receives periodically shared data from external sources.	E1. IoT devices
E1. IoT devices	DF5. Unstructured sensor data collected from the IoT sensors containing environmental data from their surroundings is periodically sent to the edge/cloud domain for its processing.	P2. Edge/cloud computing
P2. Edge/cloud computing	DF6. Processed sensor data is returned to the smart building domain to give insights about the raw data collected from the IoT sensors when requested by the users.	E1. IoT devices
P1. Hydra	DF7. Offloading of Hydria's data and processes to the edge/cloud for processing tasks with higher computational requirements.	P2. Edge/cloud computing
P2. Edge/cloud computing	DF8. Data onloading after the processing of more complex task in the edge/cloud.	P1. Hydra
P2. Edge/cloud computing	DF9. Cloud data storage.	DS3. Edge/cloud data storage

4.2.2 Mapping DFD elements to LINDDUN threat categories

This step aims to identify all the potential privacy threats that may arise based on the scenario presented in the DFD in Figure 4.1 and the detailed data flows presented in Table 4.1. As explained in 3.2.2 each element of the DFD is vulnerable to specific risks, and the type of privacy risk it faces depend on their category. In this step, I consider general assumptions regarding how the seven privacy threat categories impact each DFD element type at the level 0. This helps us understand the interactions and relationships between these elements and whether they pose privacy risks.

First, linkability refers to the inference of a connection between a pair of elements, where E, DF, DS, P represent the potential IOI that can be linked. This category applies similarly to all DFD elements, indicating that it involves the potential for two elements to be linked in some way. Potential linkability in this scenario arises from interactions between entities or the interaction between an entity with a process. This involves the exchange of data flows that may share similar attributes or be involved in the same process. For instance, user devices interact with the IoT devices to control the various functions of the smart building, such as lighting or heating. Here the exchange of data flows involves information like environmental sensor data and user preferences that can be correlated revealing patterns or relationships. In another scenario, linkability may arise with the processing that occurs in the cloud. Occupants control the setting of their environment, and that interaction generates data that is offloaded for processing, storage, analysis and long-term monitoring. Similarly to the previous case, this information can be correlated revealing patterns or relationships.

Second, identifiability refers to the establishment of a pair between a specific data subject and an attribute associated with them. This category involves distinguishing a data subject, typically represented as an entity. Thus, identifiability of an E indicates discerning a particular entity within a set of entities. In the case of DS, DF and P, identifiability means identifying the sender, receiver, data holder or accessing subject. Building upon the previous examples and considering the scenario where users interact with IoT devices and cloud computing process, identifiability can potentially arise from the correlation between user references and their work schedule. For instance, if a user consistently adjusts the smart lighting and heating systems based on their work hours, this correlation could reveal consistent and identifiable pattens in occupants' preferences and behaviours.

Third, non-repudiation involves establishing a pair, similarly to linkability and identifiability, between a subject that cannot deny its actions and the attribute or action it is associated with. This category is predominantly associated with entities. However, the non-repudiation threat category arises from DF, DS and P. Building upon the same example as before, records of exchanges of information and the completion of processes like Hydra by petition of a user may hold a subject unable to deny their involvement in these interactions. For instance, the system can save records of a user requesting a change in room temperature settings, associating actions with a specific user. While these records can enhance energy efficiency and personalization through automation in the smart building, achieving such personalized and automated processes for individual users may also result in tracking and profiling, potentially compromising user privacy.

Fourth, detectability refers to the ability to identify the presence of an IOI. This threat relates to DF, DS and P and may associate them to specific users, providing insights into the user activities, preferences, or behaviours without accessing the actual information they contain. For instance, merely detecting the existence of DF, DS or P, even without accessing their contents, can unveil relevant information about a user. Expanding on the same example, identifying a particular user's interaction within the IoT scenario and cloud computing can reveal patterns and user behaviour. Similarly, the potential detection of Hydra's process may reveal the presence of a data subject and potential changes in the environment.

Fifth, disclosure of information expands on detectability and refers to unauthorized access to the information contained in DF, DS and P. For instance, if an attacker gains access to the data flows associated with a specific user's energy consumption patterns, they could use this information to deduce the user's daily routine and behaviours.

Sixth, unawareness is specific to E and indicates the lack of user awareness, which can restrain their ability to provide informed consent regarding the use of their personal data.

While this category is primarily related to E, other components may be involved since its occurrence arise the threat of user unawareness. For instance, in the smart building context, users may not be fully aware of the constant monitoring of the IoT devices and Hydra, collecting and transmitting data due to the lack of transparency about data collection practices. This unawareness hinders the user's ability to provide informed consent or make choices regarding the use of their data.

Seventh, non-compliance affects to the whole scenario since each element bears the responsibility of ensuring that actions align with privacy policies and the consent of data subjects.

The outcome of this step is a table that associates every DFD element to each privacy category, as presented in Table 4.2 for the level-0 DFD.

Table 4.2: Mapping privacy threats to DFD element types for level-0 DFD.

DFD Element types	**L**	**I**	**N**	**D**	**D**	**U**	**N**
Entity							
E1. IoT devices	X	X			X	X	X
E2. User devices	X	X					
E3. Service Providers						X	X
Process							
P1. Hydra				X		X	X
P2. Edge/Cloud computing	X	X	X	X	X	X	X
Data store							
DS1. Hydra data storage					X	X	X
DS2. IoT devices data storage		X			X	X	
DS3. Edge/cloud data storage	X	X			X	X	
Data flow							
DF1. Sensor data to user devices	X		X	X		X	X
DF2. Data sharing user devices	X	X	X	X	X		
DF3. Sensor data service provider	X				X	X	X
DF4. Data sharing service provider							
DF5. Sensor data to edge/cloud						X	X
DF6. Processed sensor data	X		X	X		X	X
DF7. Data/process offloading	X		X	X	X	X	X
DF8. Data onloading	X		X	X	X	X	X
DF9. Cloud data storage						X	

Several assumptions were considered during this step and are listed below.

- Data flows between inside processes and their data storage are considered secure.
- Data transmissions from the IoT device to the edge/cloud occurs periodically and is considered non-threating in terms of inferring user behaviour.
- Non-compliance threats are combined as they are not specific to one part of the system but pose a risk to the system as a whole. Therefore, no distinction between the different DFD elements will be applied to this threat.

4.2.3 Identify threat scenarios

The identification and further documentation of threats is based on the threat trees provided by LINDDUN. These trees serve as a guide to conduct an in-depth analysis of privacy threats within a realistic system. Each potential threat considered in the previous step corresponds to a specific threat category that has a corresponding threat tree pattern. These trees outline the preconditions for each threat's vulnerabilities and that can be exploited in potential privacy attack scenarios. Thus, these threat trees are reviewed to fit the system under study and, as a result, a set of tables outlining all 'X's identified in the previous step are obtained and presented in Annex A. These tables serve as a source for understanding the details of each threat category.

Ultimately, by analyzing all the branches of the threat trees and documenting the potential threats the outcome of this step comprises a collection of threat scenarios presented in the form of tables. This outcome provides a comprehensive overview of the different potential privacy threats and detailed information on the threat scenarios, including the involved DFD elements, the associated LINDDUN properties, and the assets and consequences related to the threat.

Table 4.3: Threat Identification Level-0 DFD.

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
1	Profiling an individual by observing the communications between the IoT devices and the user devices outside the smart building trust boundary.	E1 E2 DF1 DF2	Linkability	Linking through a set of attributes that can serve as a quasiidentifier like a set of locations or time can uniquely link an activity to a single user to infer user behavior. These activities can be different requests made by the users to adapt the environment conditions to their preference. In the same manner, linking the use of a service to the same user through unique time patterns based on the request from the users can lead to infer user behavior. Request from the users made at the same time every day can reveal their daily routine.	Medium
			Non-repudiation	A subject may be unable to deny having interacted with an IoT device or being associated with a location or time, as transmissions are logged as evidence of communication.	
			Detectability	Detecting the existence of an item of interest can be possible by observing a communication between the IoT devices and the user devices, as well as to infer the presence of an individual and their behavior based on time patterns.	
2	Profiling and identifying a user by observing data or characteristics in the communication to the IoT devices from the user devices.	E1 E2 DF2	Identifiability	Linking through identifiers like user IDs, which are unique within the system, can be used to associate different requests of the same user and can potentially identify the individual posing the request.	High
			Disclosure of information	Sensitive information can be accessed during the communication between the user devices and the IoT devices.	

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
3	Profiling a group of individuals by gaining access to the information transmitted in the communications from the IoT devices to the service providers outside the smart building trust boundary.	DF3	Linkability	Linking through readings obtained from the structured sensor data transmitted, which apply to a household or workplace, can be used to profile a group of individuals. The correlation of these readings can derive user patterns and characterize collective behavior. For example readings from an energy consumption meter can infer the energy usage patterns of individuals within a household or workplace and it may be possible to discern daily routines. Similarly, data from occupancy sensors or access control systems can infer the movement of employees within the building.	High
			Disclosure of information	Third parties tracking and analyzing involves transferring data about the household or workplace which can lead to disclosure of sensitive information.	
4	Profiling an individual by observing the communication between the IoT devices and the edge/cloud domain outside the smart building trust boundary.	DF6	Linkability	Linking a request to an external service like the edge/cloud computing process may infer the existence of an ongoing session with a user. In the same manner, linking the use of a service to the same user through unique time patterns based on the request from the users can lead to infer user behavior. Request from the users made at the same time every day can reveal their daily routine.	Medium
			Non-repudiation	A subject may be unable to deny having interacted with the process or being associated with a location or time, as transmissions are logged as evidence of communication.	
			Detectability	Detecting the existence of an item of interest can be possible by observing a communication between the IoT devices and the edge/cloud domain.	

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
5	Profiling an individual by observing the communication between Hydra's process and the edge/cloud domain outside the smart building trust boundary.	DF7 DF8	Linkability	Linking to a particular external service like an edge/cloud can be inferred by observing the communication to the edge/cloud computing process. Additionally, linking the use of the edge/cloud process to the same user through unique time patterns based on the request from the users can lead to infer user behavior.	Medium
			Non-repudiation	A subject may be unable to deny having interacted with the process or being associated with a location or time, as transmissions are logged as evidence of communication.	
			Detectability	Detecting the existence of an item of interest can be possible by observing a communication between Hydra's process and the edge/cloud domain.	
			Disclosure of information	Sensitive data can be accessed during the communication between Hydra's process and the edge/cloud processing.	
6	Profiling a group of individuals by observing the process and storage in the edge/cloud domain derived from a user request.	P2 DS3	Linkability	Linking an activity at the edge/cloud domain with a time of day, identifier associated by the user request, or processing time to infer user behavior. Additionally, linking particular entries in the data storage to a particular household or workspace can lead to profiling a group of individuals.	Low

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
7	Identifying an individual by gaining access to the data collected in the IoT devices, shared with the edge/cloud, used for its processing and their corresponding storage.	E1 DS2 DF5 P2 DS3	Linkability	Identifying a user may be possible through gaining access to revealing attributes such as location and personal data gathered in the IoT devices, shared in the communication, and processed in the edge/cloud as well as metadata from the communication.	High
			Disclosure of information	Disclosure of real-time measurements rather than the aggregated consumption with sensitive information such as location data and time.	
8	The completion of Hydra's process at the edge/cloud can be observed.	P1 P2	Non-repudiation	A subject may be unable to deny having engaged in the process because of the registered completion of the process, as action side-effects can cause an action to be attributable to an individual.	Low
			Detectability	Detecting through side-effects of an action in the system like the processing of Hydra in the edge/cloud domain can infer the existence of an item of interest or infer user behavior.	
9	A data subject may be insufficiently aware and may be unable to set appropriate preferences on which data is collected, shared, and stored.	E1 DF1 DF5 DF6 DS2	Unawareness	The lack of awareness among data subjects regarding the collection, transmission, and storage of sensitive information about their surroundings leaves them unable to set preferences on how the data is managed or provide consent for its processing.	Medium

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
10	A data subject may be insufficiently aware and may be unable to set appropriate preferences on which data is shared with third parties.	DF3 E3	Unawareness	Data subjects may be insufficiently aware of the sharing of sensitive information about their surroundings with third parties.	High
11	A data subject may be unaware of the observation, processing, and storage of sensitive information by Hydra's process.	P1 DF7 DF8 DS1	Unawareness	The lack of awareness among data subjects regarding the observation, processing, and storage of sensitive information about their surroundings and personal images or features leaves them unable to set preferences on how the data is managed or provide consent for its processing.	High
12	A data subject may be unaware of the observation, processing, and storage of sensitive information by edge/-cloud assets.	P2 DF9 DS3	Unawareness	The lack of awareness among data subjects regarding the observation, processing, and storage of sensitive information about their surroundings leaves them unable to set preferences on how the data is managed or provide consent for its processing.	Medium
13	Data management may not be compliant with legislation.	E1 E3 P1 P2 DS1 DF1 DF3 DF5 DF6 DF7 DF8	Non-compliance	The collection, storage, processing, transmission of PII is not compliant with legislation, regulation, and/or policy.	Medium

4.3 Level 1 DFD

Secondly, I present the level-1 DFD design. This representation illustrates the process of Hydra as its decomposition into multiple subprocesses that represent the main functions of the system. It corresponds to a detailed representation of the internal processes of Hydra, helping in the understanding of its working.

The goal of this section is to provide a clearer vision of the privacy challenges that arise from the process of Hydra when integrating it into the context of a smart building. As stated for the level-0 DFD, examining privacy implications from emerging technologies informs early risk mitigation and responsible progress.

4.3.1 Define Data Flow Diagram

Extending on the process of Hydra presented in the level-0 DFD, a level-1 DFD is presented to offer a more comprehensive understanding of its internal processes in Figure 4.2. This representation includes detailed information based on Hydra's architecture 2.2.3.

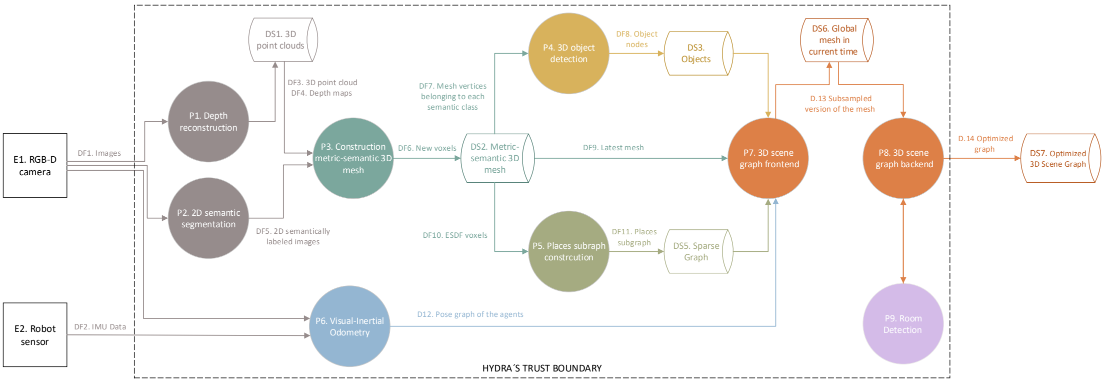

E1. RGB-D camera
E2. Robot sensor
DF1. Images
DF2. IMU Data
DS1. 3D point clouds
P1. Depth reconstruction
P2. 2D semantic segmentation
DF3. 3D point cloud
DF4. Depth maps
DF5. 2D semantically labeled images
P3. Construction metric-semantic 3D mesh
DF6. New voxels
DS2. Metric-semantic 3D mesh
DF7. Mesh vertices belonging to each semantic class
DF10. ESDF voxels
P4. 3D object detection
DF8. Object nodes
DS3. Objects
P5. Places subraph constrcution
DF11. Places subgraph
DS5. Sparse Graph
P6. Visual-Inertial Odometry
D12. Pose graph of the agents
DF9. Latest mesh
DS6. Global mesh in current time
D.13 Subsampled version of the mesh
P7. 3D scene graph frontend
P8. 3D scene graph backend
D.14 Optimized graph
DS7. Optimized 3D Scene Graph
P9. Room Detection
HYDRA´S TRUST BOUNDARY

Figure 4.2: Level 1 Data Flow Diagram.

Hydra is implemented as a highly parallelized architecture that combines low-level, mid-level and high-level perception processes to build a real-time 3D scene graph as described in Section 2.2.3. The system includes algorithms for constructing the different layers of a scene graph, metric-semantic 3D mesh, objects and agents, places, and rooms all of them gradually converging into a single node belonging to a building. The exchange of information between these distinct processes is illustrated in Figure 4.2 and explained in Table 4.4.

4.3.2 Mapping DFD elements to LINDDUN threat categories

This step aims to identify all the potential privacy threats that may arise based on the scenario presented in the DFD in Figure 4.2 and the detailed data flows presented in Table 4.4. As explained in 3.2.2 each element of the DFD is vulnerable to specific risks, and the type of privacy risk it faces depend on their category. In this step, I consider general assumptions regarding how the seven privacy threat categories impact each DFD element type at the level 1. This helps us understand the interactions and relationships between

Table 4.4: Description of data flows between entities in the level-0 DFD.

Source	Data Flow Description	Destination
E1. RGB-D camera	DF1. Real-time images capturing the environment.	P1. Depth construction P2. 2D semantic segmentation P6. Visual-Inertial Odometry
E2. Robot's sensor	DF2. IMU data describing the robot's orientation, velocity, and position.	P6. Visual-Inertial Odometry
P1. Depth reconstruction	DF3. 3D point clouds representing the surfaces in the environment.	DS1. 3D point clouds
DS1. 3D point clouds	DF4. Depth maps, in which each pixel represents the distance from the camera to the nearest surface.	P3. Construction metric-semantic 3D mesh
P2. 2D semantic segmentation.	DF5. 2D semantically labeled images, in which each pixel is labeled with a specific class to distinguish surfaces.	P3. Construction metric-semantic 3D mesh
P3. Construction metric-semantic 3D mesh	DF6. Voxels are three-dimensional pixels that represent the smallest units of a volume and they are combined to represent different elements in the environment.	DS2. Metric-semantic 3D mesh
DS2. Metric-semantic 3D mesh	DF7. Mesh vertices belonging to each semantic class are transmitted as a set based on their semantic label. The mesh is partitioned and used to classify the objects present in the scene.	P4. 3D object detection
P4. 3D object detection	DF8. Object nodes refer to objects present in the environment and contain information about their position and characteristics.	DS3. Objects P7. 3D scene graph frontend
DS2. Metric-semantic 3D mesh	DF9. Latest 3D mesh that represents the environment combining geometric and semantic information.	P7. 3D scene graph frontend
DS2. Metric-semantic 3D mesh	DF10. ESDF voxels storing the Euclidean distance to the nearest surface.	P5. Places subgraph construction
P5. Places subgraph construction	DF11. Places subgraph as a collection of nodes representing the layout of the environment.	DS5. Sparse graph P7. 3D scene graph frontend
P6. Visual-Inertial Odometry	DF12. Pose graph of the agents that contain both odometry and loop closures edges, allowing for efficient loop closure detection.	P7. 3D scene graph frontend
P7. 3D scene graph frontend	DF13. Subsampled version of the mesh that represent an initial estimate of the 3D scene graph uncorrected for drift.	DS6. Global mesh in current time P8. 3D scene graph backend
P8. 3D scene graph backend	D14. Optimized graph after the loop closure.	DS7. Optimized 3D scene graph

these elements and whether they pose privacy risks. Similarly to the level-0 DFD, each threat category explores the potential threats arising in each DFD element.

First, linkability arises from the connection between a pair of elements. In this context, I have applied this category to the potential linkage of different elements involved in the construction of the 3DSG at different levels. Initially, the incoming images from the RGB-D cameras are necessary to extract the necessary information required to construct the 3DSG. For instance, associating images or frames from cameras with specific objects or locations in the visual scene can be done by analysing visual features extracted from images, including keypoints or descriptors. Keypoints represent distinctive points or locations in an image that signify unique features like corners or edges. Descriptors are specific characteristics of these points that allow the recognition of similar points and their proper alignment. Then, associating 3D point clouds can be performed by point cloud segmentation. In Hydra, 3D point clouds may be stored both locally to provide real-time responses and in the cloud to a persistent representation. For example, by analysing the spatial distribution and properties of 3D points, like coordinates or colour, the system can infer that a cluster of points likely represent a specific objects, structures or even specific locations. Escalating the level of abstraction to higher levels, linking can be performed directly to object or places instances that share common attributes represented in the global mesh. For instance, objects that belong to the same semantic class can be grouped and counted, revealing for example how many employees work in a particular office. Similarly, agents can be linked based on objects and spaces they interact with. By analysing the interactions between agents and the various objects within a visual scene, patterns of behaviour can be established. For instance, if multiple individuals are consistently observed interacting with the same set of objects or spaces, it may suggest a shared workspace or collaborative activity.

Second, identifiability arises when there is a connection between a specific data subject and an attribute associated with them. This category can result from various ways. One way is through the direct capture of an IOI that can directly identify an individual, like a personal ID. Another way is by linking multiple IOIs that belong to the same individual, which can reduce the anonymity set. For example, in an office environment, a worker may leave behind personal items such as a phone, laptop, and even clothing in a specific area. When these objects are linked together, it becomes possible to identify that specific user. Furthermore, the correlation of valuable information collected from the environment with data from other sources like social media can also contribute to the identifiability process.

Third, non-repudiation establishes a link between an action and a subject who is unable to deny their involvement. In the context of Hydra, this category arises when a data subject can be located in different areas of the environment and a link could be established between them and different locations or objects. For example, this can be illustrated through a scenario in an office space where an important document has disappeared. If the system is capable of locating a specific user within its surroundings, that employee cannot later deny their involvement.

Fourth, detectability that refers to the ability to identify the presence of an IOI. In this context, updating the 3DSG involves updating Hydra's processes and as a consequence the final graph is renewed with the new information. For instance, when the final graph is updated, it means that changes in the environment have been detected and incorporated to the graph. These changes can include the movement or presence of objects or individuals, as well as alterations in the environment itself. Even without knowing the specific details of the changes, the act of updating the graph implies that something has occurred, which could include the presence of individuals in the scene.

Fifth, disclosure of information refers to accessing confidential data. In the context of Hydra, this category can be examined in relation to both its inputs and outputs. For instance, if an attacker gains access to the final graph representation, they have the ability to access potentially sensitive information or tamper with the graph's data. For example, an authorized individual with malicious intentions who gains access to the final graph can potentially disclose sensitive information, such as the real-time location of other individuals or their work patterns.

Lastly, the categories of unawareness and non-compliance are essentially the same as presented for the level 0 DFD. In this context, the category of unawareness states that users may not be fully aware of the constant monitoring by Hydra, which involves the constant collection and transmission of data. This lack of transparency about data collection practices hinders the user's ability to provide informed consent or make choices regarding the use of their data. Similarly, the category of non-compliance affects the entire scenario since each element bears the responsibility of ensuring that actions align with privacy policies and the consent of data subjects.

The outcome of this step is presented in Table 4.5, where every DFD element to each privacy category for the level-1 DFD.

Several assumptions were considered during this step and are listed below.

- Process threats are combined and examined as one as internal process are only susceptible to insider threat and the threats apply to all of them.
- Data flows and data stores are not considered secure since data exchange with the edge/cloud domain may occur as presented in the level 0 DFD.
- Non-repudiation and detectability threats of the processes are considered in the level 0 DFD since they pose the same threat. This threat is documented as Threat 8 in Table 4.2.3.
- Unawareness threats of the processes, data flows and data stores are considered in the level 0 DFD since they pose the same threat. This threat is documented as Threat 11 in Table 4.2.3.
- Non-compliance threats are combined as they are not specific to one part of the system but pose a risk to the system as a whole. Therefore, no distinction between the different DFD elements will be applied to this threat.

4.3.3 Identify threat scenarios

By following the same procedure as with the level-0 DFD, threats were identified by referencing the threat trees provided by LINDDUN. This process results in a set of tables that outline all the 'X's identified in the previous step and that are presented in Annex A. Analogously to the level-0 DFD, the outcome of this step consists of a collection of threat scenarios presented in the form of tables, providing a comprehensive overview of the different potential privacy threats.

Additionally, in this section the prioritization of the identified threats has been included. The prioritization has been carried out following a qualitative risk assessment, where the values used for the prioritization are non-numerical values. This straightforward approach measures the likelihood and the impact of a threat scenario with values low, medium, and high.

Table 4.5: Mapping privacy threats to DFD element types for levle-1 DFD.

DFD Element types	**L**	**I**	**N**	**N**	**D**	**U**	**N**
Entity							
E1. RGB-D camera						X	X
E2. Robot's sensor						X	X
Process							
P1. Depth reconstruction			X	X		X	X
P2. 2D semantic segmentation			X	X		X	X
P3. Metric-semantic 3D mesh construction			X	X		X	X
P4. 3D object detection			X	X		X	X
P5. Places subgraph construction			X	X		X	X
P6. Visual-Inertial Odometry			X	X		X	X
P7. 3D scene graph frontend			X	X		X	X
P8. Room detection			X	X		X	X
Data store							
DS1. 3D point clouds	X	X	X			X	X
DS2. Metric-semantic 3D mesh						X	X
DS3. Objects	X					X	X
DS4. Agent nodes						X	X
DS5. Sparse Graph	X					X	X
DS6. Global mesh in current time	X	X	X			X	X
DS7. Optimized 3D Scene Graph	X	X	X			X	X
Data flow							
DF1. Images	X	X	X	X	X	X	X
DF2. IMU data	X		X	X		X	X
DF3. 3D point clouds	X	X	X			X	X
DF4. Depth maps						X	X
DF5. 2D semantically labelled images						X	X
DF6. New voxels						X	X
DF7. Mesh vertices from each semantic class						X	X
DF8. Object nodes	X					X	X
DF9. Latest mesh						X	X
DF10. ESDF voxels						X	X
DF11. Places subgraph	X					X	X
DF12. Pose graph of the agents						X	X
DF13. Subsampled version of the mesh	X	X	X			X	X
DF14. Optimized 3D scene graph	X	X	X	X	X	X	X

Table 4.6: Threat Identification Level-1 DFD.

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
1	Images and features of the environment can be obtained and linked through combination or analysis of data.	DF1 DF3 DS1	Linkability Disclosure of information	A link between similar images and features of the environment can be inferred even without knowing the source or context of the images. This information could be used to recognize patterns in the environment and to identify the location where the images were taken. Additionally, reconstruction of detailed comprehensive images of the scene can be obtained from 3D point clouds and their associated attributes even if the source images are discarded. The reconstructed images can be used as source images and pose the same threats. [55].	High
2	Metadata from real-time images and IMU data, such as coordinates and time stamps, can reveal the exact location and time where the image was taken.	DF1 DF2 DF3 DS1	Linkability	A link between a set of locations or timestamps can be used to uniquely link activity to a single user and infer user behaviour. Additionally, reconstruction of detailed comprehensive images of the scene can be obtained from 3D point clouds and their associated attributes even if the source images are discarded. The reconstructed images can be used as source images and pose the same threats [55].	Medium
			Non-repudiation	A data subject may be unable to deny their presence or being associated with a location at a certain time from the metadata contained in the images and IMU data.	

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
3	Objects represented in the scene graph can be observed and linked based on properties.	DS3 DF8	Linkability	A link between different objects present in the scene can infer the quantity of objects of a same semantic class which can be used to profile a group of individuals.	Medium
4	The topology of the graph can be inferred from the sparse graph, which preserves the underlying structure of the scene.	DS5 DF11	Linkability	A link between the vertex of the graph can infer its topology by obtaining the free-space paths and therefore the layout of the scene.	Medium
5	The 3D mesh in real-time can be observed and a data subject can be linked to a certain object or location, inferring subject behaviour.	DS6 DF13 DS7 DF14	Linkability Non-repudiation	A link between vertex from the sparse scene graph or rooms from the optimized scene graph can be inferred based on locations that an agent node has visited. Linking rooms may infer user behaviour based on their movements within the building and may hold a data subject unable to deny having been associated with a certain location. A link between agent nodes can be inferred from agent nodes that have visited the same location simultaneously or over a period of time. Linking agents may infer user behaviour and may render subjects unable to deny having engaged in a conversation or in a certain activity. A link between objects and agent nodes can be inferred based on the interactions of agent nodes with those objects. A subject may be unable to deny having interacted with those objects.	Medium

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
6	Captured images of the environment can be used for purposes such as face recognition that can be used to identify a data subject [56].	DF1 DF3 DS1	Identifiability	Face recognition processing can be applied to the images deriving certain attributes that can be combined to identify an individual. Additionally, images containing PII, such as credit card number or social security number can lead to uniquely identifying a data subject. Additionally, reconstruction of detailed comprehensive images of the scene can be obtained from 3D point clouds and their associated attributes even if the source images are discarded. The reconstructed images can be used as source images and pose the same threats.	High
			Non-repudiation	A subject may be unable to deny their presence or being associated with a location from the captured images.	
7	Objects that belong to the same individual are represented in the scene graph and can be observed and linked to reveal their identity.	DS6 DS7	Linkability	A link between different objects present in the scene can infer the presence of a data subject, even without revealing their identity.	High
			Identifiability	Furthermore, the combination of objects that contain unique references to the identity of the data subject can reveal their identity.	
			Non-repudiation	A subject may be unable to deny having been associated with a specific location.	
8	Observed communications to or from external entities may infer the exitance of relevant information.	DF1 DF2 DF14	Detectability	The existence of an item of interest may be inferred from observing the communciation to or from an external entity, even without disclosing the transmitted data. New images and the generation of the 3DSG in real-time may disclose the presence of a subject in the scene.	Medium

Threat	Summary	DFD elements involved	LINDDUN properties	Assets involved and consequences	Priority
9	Gaining access to the communications to or from external entities may cause the disclosure of sensitive information.	DF1 DF14	Disclosure of information	Sensitive information such real-time images and the 3DSG containing the real-time representation of the environment may be disclosed.	Medium
10	A data subject may be insufficiently aware and may be unable to set appropriate preferences on which data is collected from the external devices.	E1 DF1 DF5 DF6 DS2	Unawareness	The lack of awareness among data subjects regarding the collection of sensitive information about their surroundings and features leaves them unable to set preferences on how the data is managed or provide consent of its processing.	Medium
11	Data management may not be compliant with legislation.	Complete system	Non-compliance	The collection, storage, processing, transmission, of PII is not compliant with legislation, regulation, and/or policy.	Medium

5. Solution Space

Following the previous chapter, the last three steps presented in Section 3.2.2 are followed to obtain the results from applying the LINDDUN privacy threat modeling to Hydra's pipeline. This chapter is organized in four sections. The first section provides a comprehensive overview of the different stages that the data go through to obtain visual representation of the environment. The following three chapters correspond to these stages, which are input protection, data protection, and output protection. At each stage, the last three steps of the LINDDUN methodology are followed. Firstly, the prioritization of the identified threats that was presented Sections 4.2 and 4.3, along with the threat identification. Following that, mitigation strategies are proposed based on the nature of the protection approaches. Finally, building upon the mitigation strategies from the previous step, appropriate PETs will be proposed for each category, and they will be associated to the respective threats for which they provide a solution.

The research of this work has been primarily dedicated to the critical analysis of existing solutions and the selection of PETs to address and mitigate the different privacy threats associated with the level-1 DFD.

5.1 Solution Design

The solution space aims to tackle the challenges at the level-1 DFD, safeguarding sensitive information and ensuring privacy protection in the context of detection, transformation and rendering of visual information for environmental mapping. For this purpose, the same approach that was made for mix reality in [57] has been followed and Hydra's pipeline can be simplified as represented in Figure 5.1. Firstly, the RGB-D camera that provides the images and the robot's sensor that provides the IMU data are responsible for sensing the real environment and gathering the information for its detection. Then, the collected data is processed, and different transformations will be applied based on the desired output. Lastly, the modelled environment is rendered and the 3DSG is shared with third parties and applications for its use.

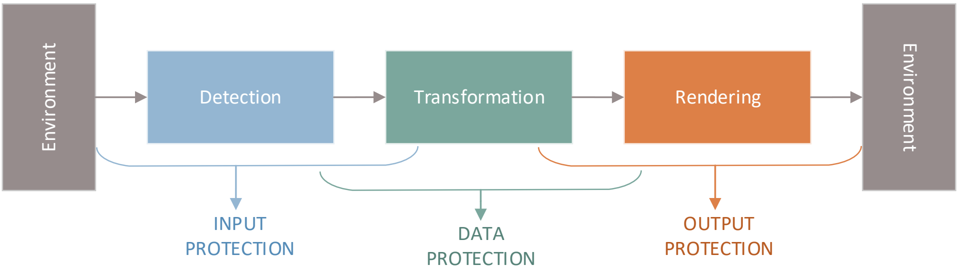


Figure 5.1: Generic data life cycle in perceptual application [57].

In the following sections, the three categories of protection will be discussed and mitigation strategies and PETs are proposed.

5.2 Input protection

The first category encompasses the raw visual data collected from the environment, which is then integrated into Hydra for further processing. Hydra's main purpose is to provide a visual representation of a scene, which requires gathering information about the environment, including construction elements such as walls and doors, and objects such as desks and chairs, necessary for an accurate depiction of the scene. However, other items present in the scene, such as personal belonging like wallets or mobile phones, which do not contribute to modeling, may unintentionally be captured. Hence, the gathered information may include user sensitive data as well as sensitive data related to other entities, such as bystanders. Targeted physical elements that assist in building the model can be considered active inputs, while untargeted objects that are not necessary for representing the environment can be interpreted as passive inputs [57][58].

5.2.1 Elicit mitigation strategies

This category can be associated with the branch of concealment of data from the mitigation techniques discussed in Section 3.2.2. This branch focuses on guarding the exposure of sensitive data, ensuring that only necessary information is shared with the system. For perceptual applications, protection approaches in the input side typically involve the use of input sanitization techniques. These measurements can be employed as an intermediate layer, situated between the sensor interfaces and the application layer, with the aim to remove latent and sensitive information from the input data stream. They can be further classified based on the type of policy enforcement they employ, whether it is intrinsic or extrinsic. Intrinsic input sanitization policies are typically user-defined, allowing users to specify the degree of sanitization that will be applied. However, intrinsic policies may potentially have a myopic view of the privacy preferences since they are user-defined. On the other hand, extrinsic input sanitization policies receive privacy preferences from the environment and can prevent the capture of sensitive objects that may not be considered by the intrinsic policies. This limitation exists because intrinsic policies can solely provide protection for inputs that are explicitly mentioned within the policies [57][58].

5.2.2 Selection of PETs

This section analyses existing solutions that could potentially mitigate the identified threats, exploring both their potential and the challenges of integrating them into Hydra's pipeline. The variety of available options highlights the difficulty of finding customized solutions. Within this field of research, the majority of works are focus on Augmented Reality (AR) and Mixed Reality (MR), resulting in the proposal of several approaches. These approaches can be mainly categorized as follows [57][58]:

- Visual Information Sanitization: Early approaches mainly revolved around sanitizing visual media to remove sensitive content. Various methods have been employed within this category. Some apply intrinsic policies, like Darkly [59], that restricts the access to complete information, or the context-based intrinsic sanitization framework [60]. Others apply extrinsic policy enforcement, enabling objects and bystanders to communicate privacy preferences like MarkIt [61] or Cardea [62] that build in fine-grained visual protection based on the user's preferences. More recent work like Virtual Curtains [63], focus on providing real-time policy enforcement.

- Visual Information Abstraction: Abstraction seeks to minimize direct access to raw visual feeds by providing only essential information to applications. This approach follows the concept of least privilege and has been applied to secure gesture detection and spatial information in MR environments like Recognizers [64], Prepose [65], or SafeMR [66] that introduces an object-level abstraction system designed for MR applications.

Each of these approaches addresses the challenge of protecting visual information in MR contexts from a different angle. Their respective strengths and limitations call for further exploration and evaluation to determine their suitability and effectiveness in different scenarios. In this study, I have selected several of these technologies to evaluate their potential applicability in the Hydra scenario.

DARKLY

The Darkly system [59] proposes a multi-level feature sanitization to address some of the key challenges faced by perceptual applications in the input stream. These challenges include the recognition and isolation of objects, the discerning of which inputs should be revealed and the appropriate form to disclose these inputs. Darkly employs mechanism such as access control to manage the access to visual inputs, algorithm transformations to obfuscate the visual data and user audit to enable transparent tracking of actions carried out on the visual inputs. For these purposes, it relies on OpenCV [67], an open-source library including computer vision algorithms for image processing, video analysis, 3D reconstruction and object detection among other functionalities.

Darkly mainly focusses on two approaches, opaque references and declassifier functions. The first approach is based on replacing the raw visual inputs with opaque references before its transmission. Opaque references are identifiers that allow interactions with the referenced entity without exposing its content. The aim is to limit the access to the raw input stream and to only share the necessary information with untrusted third parties and applications. Pointers to image data are replaced with opaque references, which allow applications to operate on these inputs without directly accessing them. Thus, applications cannot dereference opaque references, but they can be passed to and from OpenCV functions, which then operate on true perceptual data. To differentiate between opaque references and real pointers, a threshold is set, stablishing that all valid pointers must be greater than a value and the opaque references are below that value [59].

The second approach involves applying privacy transforms to features or object in the input stream before releasing the data to third parties or applications. It is mainly intended for applications that require access to specific features of the data while ensuring that sensitive information is not exposed. For this purpose, declassifier functions are used to reduce or eliminate sensitive data from the input stream while still preserving the desired features. In this approach, users are able to determine the different levels of transformation, defining the amount of detail that is provided. This is done in the Darkly console window that displays a visual representation of the features, which enables the adjustment of the degree of transformation to be applied. These transformations are specific to the declassifier and application independent. Higher values of transformation entail more abstract and generic representations, while lower values of transformation remove fewer details as shown in Figure 5.2. Transformation techniques involve sketching and generalization.

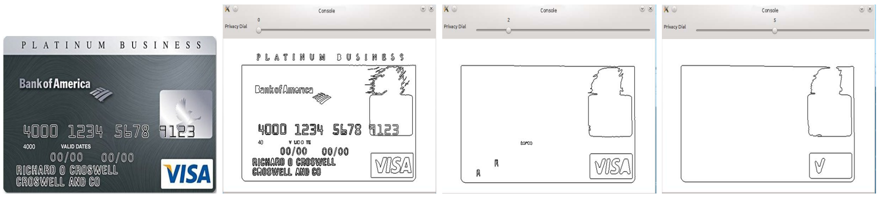


Figure 5.2: Example of the possible outputs of applying sketching transform on a credit card image [59].

Sketching is founded on identifying the contours of the image and the use of two different low pass filters. Initially, a low pass filter is used to blur the image and remove the small-scale details that may contain contextual information like credit card details. This step preserves the large-scale features to continue with the contour detection. Contours represent the transitions between different regions or objects within an image and they refer to the points with fixed greyscale color values. In a grayscale image, each pixel has a specific intensity value that corresponds to the brightness level, ranging from black that is represents the lowest intensity, to white representing the highest intensity. Thus, contours are pixels that remain constant in their brightness and that match a specific threshold. The last step involves applying another low pass filter to the counted image in order to remove contours that contain excessive entropy. Entropy in this context refers to the level of complexity of a given image, being more diverse and unpredictable with higher entropy and more uniform and repetitive with lower entropy. The combination of these three steps guarantee that details of the image remain undisclosed after the transformation and prevents countermeasures such as image deblurring [59].

On the other hand, generalization aims to provide a generic representation from a predefined collection of samples. It involves capturing the essential characteristics or common features of objects, faces or other entities and derive a more generalized representation. Darkly proposes a generalization-based privacy transform referred to as cluster-morph, which is a complex approach to generalization. Using face generalization as an example, this approach substitutes the original face with a generic face image by employing an algorithm to generalize the facial features and selects an image from a predefined collection of programmatically generated face images, resembling the way avatars are made. As contrary to other approaches, Darkly's intention is not achieving k-anonymity, but it aims to identify a canonical representation based on a globally predefined dataset [59].

Context-based intrinsic sanitization framework

Furthermore, Darkly is improved by a context-based intrinsic sanitization framework [60] that builds upon its non-contextual policies. The key advancement in this framework is the ability to determine the presence of sensitive entities like subjects, objects, and locations in the captured images and automatically apply sanitization. Thus, sensitive features undergo a process of sanitization by being blurred out, ensuring that information is concealed, and sensitive locations are entirely deleted. This approach is implemented by five main modules that rely on the contextual information obtained from the sensors as shown in Figure 5.3.

The first modules, Human activity recognition (HAR) and ambient environment detection (AED), are used to evaluate the context and identify potential sensitive subjects. HAR aims to detect and classify different user activities by analyzing accelerometer data. AED, on the other hand, focuses on determining the environmental context, distinguishing between indoor and outdoor settings. Then, the image classifier module (ICM) leverages on the labels received from the HAR and AED modules and determines the appropriate sensitive subject's detections module (SSDs) that needs to be applied. SSDs are responsible for identifying sensitive subjects relevant to each specific context, which includes the user activity and the environment. This allows further processing by the policy enforcement module (PEM), either blurring or deleting the sensitive subjects, according to the specific requirements and policies. The policies applied by this framework are user-defined, reason for which this approach is considered as intrinsic policies in spite of the context-based nature of the sanitization [60].

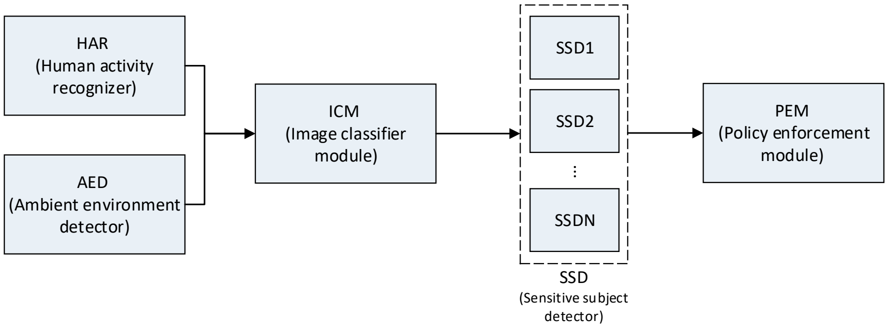

HAR
(Human activity recognizer)
AED
(Ambient environment detector)
ICM
(Image classifier module)
SSD1
SSD2
⋮
SSDN
SSD
(Sensitive subject detector)
PEM
(Policy enforcement module)

Figure 5.3: General architecture and modules of context-based intrinsic framework [60].

Virtual curtain

Recent work in input sanitization focuses on developing a novel privacy control framework for Augmented Reality (AR) systems. The virtual curtains approach [63] addresses the privacy concerns that arise from continuous-sensing cameras and aims to support two functions. Firstly, users should be properly informed regarding the specific directions from which privacy-sensitive objects should be blocked. Secondly, users should be able to implement fine-grained control over the input policies, allowing them to define their preferences and requirements for privacy control. This framework is implemented in three modules that unitedly transform users input into control policies and act in real-time on those policies, ensuring that sensitive information is removed before applications gain access the input stream [63].

The first module, policy configuration, enables users to accurately adjust the size and position of the virtual curtains around the desired objects, granting them intuitive control over the blockage of the targeted physical objects, as shown in Figure 5.4. This figure represents three possibilities to apply virtual curtains to different objects. The first shows a one-sided virtual curtains, which can be used for flat objects or surfaces, while the other two show a three and four-sided virtual curtain for volumetric objects. Users can define the size and position of the virtual curtains via the buttons in the application interface, which are located in the lower right part of the figure. Additionally, users can perform actions such as selecting, translating, scaling and rotation through a gesture recognition function to operate on the virtual curtains. This intuitive process allows users to adapt the virtual curtains to the desired objects [63].

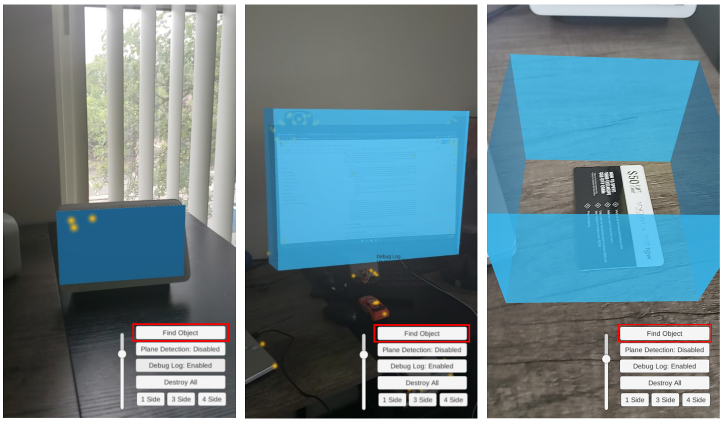


Figure 5.4: Examples of one-side, three-side and four-side virtual curtains applied to 3D objects [63].

The second module, policy capture and registration, transforms users' needs into input control policies. Firstly, the policy capture process associated the target object with the virtual curtain configured in the previous module. This association is achieved by identifying the target object that overlaps with the 2D frame projections of the virtual curtains. It uses a CNN model that detects the objects present in a frame and returns their object class, bounding box, and a confidence score for each detection. This model then finds the object whose bounding box's center is closest to the center of the virtual curtains and determine the Binary Robust Invariant Scalable Keypoints (BRISK). Secondly, the policy registration module defines the policies to be implemented on each object, which are then stored to make them accessible for the next module. This policy level of detail allow to uniquely identify objects. The definition of the policy is based on the type of object, the BRISK features, and operation to be performed on each object. These operations such as blurring, pixelating, adding filters or replacing with photos, allow access to the features of the blocked objects, while protecting sensitive information, as shown in Figure 5.5. This figure shows an example of applying the virtual curtains to a computer screen. In this case, the virtual curtain placed in front of the screen effectively blurs it from the front view, as illustrated in the second picture from the left. However, this configuration is not effective when viewed from the side, as shown in the other two images. In such cases, the virtual curtain is not positioned to block the laptop from a side view due to a misalignment between the bounding box center and the object's center, which result in a leakage of information [63].

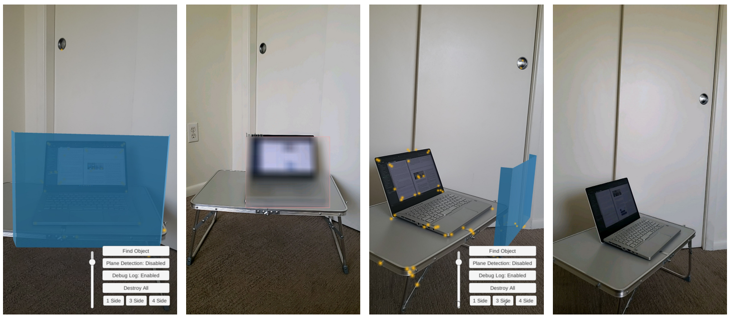


Figure 5.5: Example where successful blocking from the front side with virtual curtains effectively secures the laptop, while an unsuccessful blockage, due to misplaced virtual curtains, results in a leak from the side view applied to the laptop [63].

The third module, the policy enforcement, protects in real-time the input stream by applying the previously defined policies. This module detects the object the user intends to block by making use of a trained machine learning-based model that is loaded into the application during its initialization. The policies are loaded and the camera frames are accessed to be sent to the machine learning model. This model is able to detect objects and then, by using an object tracking technique, the objects are recognized in all the frames. Then, the identified objects are altered across all frames accordingly to the policies and subsequently shared [63].

5.2.3 Input protection applied to Hydra

During the input stage, the primary threats that arise from perceptual applications like Hydra result from the overcollection and aggregation of raw visual data in time and space, which may contain sensitive information. To recapitulate, the threats presented in Section 4.3 and listed in Table 4.3.3 as Threat 1, Threat 3, Threat 6, Threat 7 and Threat 9 originate from the collection of multiple images over time and space from a scene. The overcollection of these images, which may contain sensitive information like personal objects (e.g., as credit cards) or biometric data (e.g., facial features that uniquely identify an individual), as well as sensitive locations like bathrooms, poses a severe risk of privacy. Additionally, the privacy threat of unawareness included in Table 4.3.3 as Threat 10 originate from the lack of consciousness about a particular situation. These risks can be mitigated by the previously discussed frameworks, which although primarily focused on addressing risks in AR, can also be applied to spatial perceptual applications like Hydra. In general terms the most common solution is the least-privilege principle that conceals sensitive information.

Darkly presents two main approaches: opaque references and privacy transforms. The use of opaque references ensures that no sensitive information is shared with external entities, limiting the access to the raw input stream while still allowing interaction with these entities. On the other hand, some cases require access to specific features and applying privacy transforms allows to share the required information without disclosing further details. These techniques enable the identification and interaction with elements without disclosing all of their features. Despite Darkly being a mature approach, its principles set the basis for protecting sensitive information, which still maintains its relevance. The main drawback of this framework is that it has not been applied in real-time scenarios and lack immediacy, which presents challenges when attempting to apply it to frameworks like Hydra. Furthermore, the later improvement allows to automatically apply sanitization after detecting the presence of sensitive entities, allowing to blur sensitive features and to completely erase sensitive locations. This improved framework was tested in both outdoor and indoor environments. In indoor scenarios, this approach effectively blurred faces and screens in most of the cases. On the contrary, it was not able to detect smaller objects like credit cards. When dealing with entire sensitive images, such as bathrooms or toilets, this approach blurred the entire frame. Overall, the results presented in this approach achieved a success rate of approximately 62% in indoor scenarios. This result represents an average that takes into account static situations, with a success rate of 71.5% when individuals are sitting and 57% when they are standing, as well as moving situations, which had a success rate of 58%. This framework solves the main issue found in Darkly and has been tested in real-time scenarios. However, the immediacy that frameworks like Hydra require will still pose a challenge with the success rates achieved by this framework.

Virtual curtains focus on covering the targeted physical objects in real-time while assuring the users' awareness of how their personal data is being handled. This recent work addresses some of the limitations of the previous approaches by providing users with the ability to define their privacy preferences in finer detail. This level of control empowers users to apply virtual curtains for privacy protection and translate those settings into actionable privacy policies. It allows users to make more precise adjustments, effectively overcoming the restrictions in Darkly that limit the customization of privacy policies. This approach has been tested in real-time to determine its effectiveness in concealing objects from the input stream while ensuring alignment with the user expectations. For these purposes, they evaluate static and dynamic scenarios where the user has visual access to a laptop from various angles: front, left and top and the freedom of movement with the capturing device. The results show that the system effectively captures policies in static scenarios but exhibits slightly reduced precision in dynamic situations. This lower precision can be attributed to two factors: user movement causing a subtle shift in the virtual curtain, resulting in misconfigured policies, and the rapid device movement leading to blurred frames and failures in object detection. Additionally, frames that only contain portions of an object also pose issues when less than 30% of the object is captured in the frame. This work also highlights the computational demands of real-time object detection, where compensatory measures like multi-threading and object tracking exhibit limited performance due to resource constraints on the devices, prompting the exploration for future works. Advancements like virtual curtains, which aim to enhance the user experience within frameworks like Hydra, offer a promising solution for improving privacy. The level of control provided by this approach aligns with the evolving landscape of privacy concerns in scenarios that require immediacy and real-time responses.

5.3 Data protection

After the data has been collected, it undergoes processing to generate an output. However, once the data has been gathered, user have no control over how their data is being handled and shared with third-party agents, which can access personal information. The aim is to leverage this information without disclosing any user sensitive information through the data life cycle. Consequently, data protection measures can be categorized at different stages, during aggregation, during processing and during storage. Although data aggregation can be considered part of input data protection, a different approach can be taken as how data is managed after it has been sensed [57][58].

5.3.1 Elicit mitigation strategies

This category can be associated with the branch of guarding association from the mitigation techniques discussed in Section 3.2.2. This branch focuses on controlling the associations after disclosure and minimizing their exposure. Perceptual applications rely on the overcollection and aggregation of extensive volumes of data for its processing and deliver real-time outputs, which can be mitigated by data minimization. Other protection approaches in this stage typically involve the use of k-anonymity and differential privacy. K-anonymity focus on guarding the individual identities from a bigger group that share the same quasi-identifiers, belonging to the same anonymity set. This technique involves data perturbation or manipulation to ensure that each record remains indistinguishable from at least k-1 other records. The drawback of this approach is that is suffers from scaling problems when having high number of sensors or input data sources, making it difficult in be implemented in this kind of scenarios. Differential privacy relies on adding noise or inserting randomness to data to provide plausible deniability and unlinkability. Additionally, protected data storage solutions like personal data stores (PDS) provide a secure way for storing user data. PDS allows users to manage access to their information, determining which applications can use it, as well as monitoring the data flows in and out. This can be achieved by implementing a sand-box mechanism that monitors the data that is provided to the applications [57][58].

5.3.2 Selection of PETs

Similarly to the input protection, there is a multitude of possible approaches under consideration to mitigate the challenges of protecting data during processing. Each of these approaches tackles the issue from a distinct perspective, and it is essential to examine their individual strengths and limitations. Some of the techniques that can be employed in this stage can be categorized as follows [57][58]:

- Encryption-based techniques such as homomorphic encryption, allow computations on encrypted data. For instance, HE-SIFT [68] performs bit-reversing and local encryption on raw images before feature description and leveled-HE [69] that aims to improve the computation time of the precious approach.
- Secret sharing methods involve splitting data among untrusted parties, allowing computations without revealing individual data. For instance, secure multi-party computation enables data processing from multiple sources without disclosing sensitive information. SECSIFT [70] splits the SIFT features among a set of cloud servers that also improves computation time of HE-SIFT.
- Virtual reconstruction approaches leverage the inherent artificial nature of MR to provide sanitized virtual reconstructions of physical spaces instead of sharing complete 3D data. This approach balances sanitization and utility. For instance, one approach involves the use of 3D line clouds instead of the traditional 3D point clouds [71], which obfuscates detailed 3D structural information. Another technique that represents surfaces as a set of planes [72], focuses on mitigating spatial inference attacks.

Each of these approaches addresses the challenge of protecting visual information in MR contexts from a different angle. Further exploration and evaluation are needed to determine their suitability and effectiveness in different scenarios, including their potential applicability in the Hydra scenario, as explored in this study.

Privacy Preserving Image-Based Localization

This framework addresses the privacy concerns associated with image-based localization that arise from the constant storage of 3D point clouds of a 3DSG. The generation of 3D maps of the environment and the determination of the camera pose estimation require the constant storage of information of the scene. Generally, the source images are eliminated once the mapping process is complete. However, the 3D point clouds obtained from these images are continuously stored. These can be used to accurately reconstruct replicas of the original images of the scene, as well as inferring the layout of the scene including the presence of any confidential objects. This approach addresses the threat of disclosure of sensitive information by replacing the traditional 3D point cloud representation with 3D line cloud representation. This innovative representation conceals the underlying geometry of the scene, as shown in Figure 5.6, while still providing enough geometric constraints to support reliable and precise estimation of the camera's positions and orientation in six degrees of freedom (6-DOF) [71].

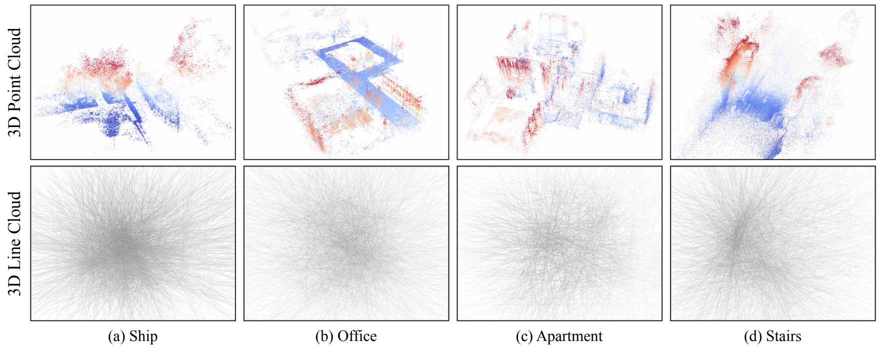


Figure 5.6: Original representations with 3D point clouds with the corresponding 3D line cloud representation [71].

To obtain this representation, known as 3D line clouds, the 3D lines and their associated features descriptors are stored, while discarding the original 3D point locations. This study begins by considering the localization of a single image as the sole input and it is extended to the more complex scenario of jointly localizing multiple images. The process of localizing an image involves determining the precise position and orientation of the camera relative to the surrounding environment using visual information from the images [71].

The traditional camera pose estimation method relies on having information about the scene's structure in the form of 3D point clouds, which discloses the scene's underlying geometry as a necessary part of its operation. In practice, these 3D point clouds are often obtained by reconstructing 3D structures from images using a technique known as structure from motion (SfM). This conventional approach involves matching features in 2D images to corresponding features in 3D point clouds. The absolute pose of a camera is described by its rotation and its translation, denoted as P = [R T], which allows to describe its movement and orientation in 3D space. In this approach, every matching between the features of 2D images and 3D points introduces a pair of geometric constraints: the depth of the observed image point and the lifted representation of these points in projective space. To estimate the 6DOF, a minimum of three 2D-3D correspondences are required. This problem is commonly referred to as the pnP problem. Following this initial estimation, it is refined through a process that seeks the most probable estimate based on the Gaussian error model for image observations [71].

Alternatively, the proposed privacy-preserving method is based on the key idea of obfuscating the scene's geometry in order to conceal its underlying geometry. This approach suggests converting each 3D point into a 3D line, which involves creating a line that has a random direction in the three-dimensional space and passes through the original 3D point, as depicted in the previous figure. This transformation changes the representation making it impossible to recover the 3D point's location due to the randomness of the line's direction. In this case, the correspondence to the projected 2D lines and 2D points reduces the geometric constraint to one making it now necessary to have at least six 2D point to

3D line correspondences to obtain the 6DOF. Furthermore, this approach is extended to jointly localize multiple cameras, which differs from the single-camera case in its parameterization. Instead of determining a separate pose for each camera, a single transformation is calculated to adjust to the entire set of cameras at once [71].

Broadly, this framework considers three different scenarios in which user privacy can be compromised. The first scenario pertains to situations where the scene itself is confidential, resulting in potential unauthorized access even when employing secure servers for cloud-based storage and processing. The second scenario includes situations where the scene is not confidential, yet it contains sensitive objects or information. The objective in such scenarios is to enable the persistent generation of 3D maps while preserving the confidentiality of sensitive information and ensuring user awareness of the capturing process. The third scenario involves sharing 3D maps among authorized users, requiring the confidential encoding of the map to protect privacy. Furthermore, the privacy-preserving approach has been tested with 15 real-world datasets of both outdoor and indoor environments, as represented in Figure 5.6. These tests aim to analyze the accuracy and robustness of this approach in estimating the absolute camera pose, while comparing it to the traditional approach. The results indicate that, in terms of accuracy and reliability, the traditional approach slightly outperforms the privacy-preserving approach, as it relies on two constraints for pose estimation compared to the one constraint that the proposed method uses. In terms of runtime, the privacy-preserving approach is slower than the traditional approach. However, it demonstrates its suitability for real-time scenarios. Additionally, these tests show that the proposed approach is robust when handling different situations. In summary, the 3D line cloud approach offers a permanent and memory-efficient transformation, resilient against privacy attacks in real-world scenarios [71].

Conservative Plane Releasing for Spatial Privacy Protection in Mixed Reality

This framework focuses on the threat of spatial inference, in which adversaries may attempt to infer the user's location or even extract information about user poses, their movements, or changes in their surroundings based on historical 3D data of their environment. 3D data offers an accurate representation of the user's environment and are constructed using a set of point clouds. A point cloud consists of a collection of oriented point and mesh information, which indicates the connections between these points to form surfaces. Each point is formed by their spatial coordinates in space {x,y,z} and a normal vector $\{n_x, n_y, n_z\}$ that indicates the orientation of the surface to which the point is associated. This works addresses two levels of inference: inter-space inference, where the adversary aims to identify the general location of the user, and intra-space inference, where the adversary seeks to determine the specific location of the user within a given space. Additionally, this approach assumes that the attacker has prior knowledge about the space and can only infer spaces that the user has previously visited [72].

The threat of inference involves two main steps, given that the adversary possesses prior knowledge about the spaces. Firstly, a reference model is created by using 3D description algorithms on the known spaces to capture their unique features. Then, this model is used to test and infer the characteristics of unknown spaces by matching their 3D descriptors to those stored in the reference model. This way, the attacker can produce a hypothesis to reveal the inter-space location and the intra-space location. This approach proposes two protection strategies to mitigate the spatial inference threat, partial releasing and planar spatial generalization. These strategies are used jointly because it has been proved that spatial generalization alone is an ineffective approach for spatial privacy [72][73].

Partial spaces are intended to minimize the amount of information shared by limiting the data released through point clouds. This approach reveals only a portion of the raw spaces, which involves selectively sharing segments of the space with varying radius, as depicted in Figure 5.7a. On the other hand, spatial generalizations involve representing any 3D surface within a space as a set of planes, as shown in Figure 5.7b. This approach can be considered a form of sanitization as surface-to-plane generalizations may unintentionally remove finer details below the desired level of abstraction. Moreover, an attacker can replicate the generalizations applied on the released point cloud data and is able to adapt their strategies accordingly. Therefore, a conservative plane releasing approach is proposed, building upon plane generalizations by restricting the number of planes generated during generalization, as shown in Figure 5.7.

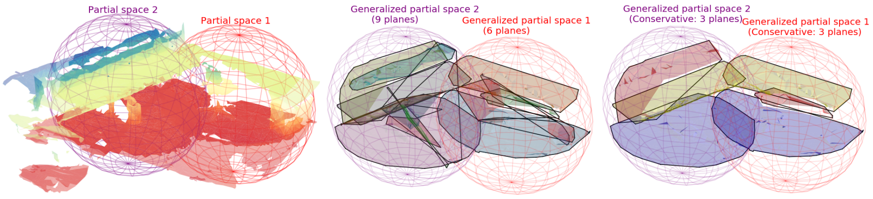

Partial space 2
Partial space 1
Generalized partial space 2
(9 planes)
Generalized partial space 1
(6 planes)
Generalized partial space 2
(Conservative: 3 planes)
Generalized partial space 1
(Conservative: 3 planes)

Figure 5.7: a) Partial spaces b) Spatial generalization c) Conservative plane releasing [72]

To prove it, this approach designs an adversary, aiming to infer the location of a user as they move around by using the gathered data and serve as a tool to test and evaluate the effectiveness of the proposed spatial privacy approach. The results obtained from this approach show that by sharing 11 generalized planes of the user's space using a radius less than 1.0 m, the attacker is not able to identify the user's exact location at least half of the time. Furthermore, up to 17 generalized planes can be revealed if the radius is minimized down to 0.5 m, while ensuring data utility and privacy preserving [72].

5.3.3 Data protection applied to Hydra

During the data processing stage, the main threats arise from the handling of vast amounts of data. Hydra relies on the overprocessing and the persistent storage of 3D point clouds for a more realistic representation of the environment, but this data can potentially reveal sensitive information. The threats outlined in Table 4.3.3 as Threat 1 and Threat 6, apart from being considered within the scope of input protection, also pose risks during the data processing stage. These threats arise due to the constant processing and storage of 3D point clouds of these images, which could be used to reconstruct replicas of the original images. Furthermore, the threat mentioned in Table 4.3.3 as Threat 4 refers to the inference of the scene's layout and the exposure of captured objects and Threat 5 compiles various potential threats associated with user activity within the scene. By tracking agent movement around the scene, a user can be linked to different spaces/rooms, objects, and other agents. These privacy threats can be mitigated by the previous discussed frameworks, which can be applied as an intermediate layer to the input point cloud before proceeding further down the pipeline. As previously mentioned, some data protection approaches intersect with input protection and most of the data sanitization and abstraction methods focus on blocking parts of the input stream and that can potentially limit the data utility. By proposing other approaches that focus on the manipulation and transformation of sensitive information there is a trade-off between data utility and privacy preserving [73].

The privacy preserving image-based localization approach aims to conceal the underlying geometry of the scene and also protect sensitive objects by proposing a representation based on 3D lines generated from the 3D point clouds. As mentioned, the continuous storage of sensitive information like 3D point clouds inherently poses a privacy risk. The act of retaining this data can potentially expose sensitive information, making it vulnerable to unauthorized access, data breaches, or misuse. This representation aims to generate a representation where no information can be inferred from looking at the 3D line clouds. It uses a simplified code to store the 3D lines, encoding their direction in just 1 byte and representing their position with 2 floats. This reduces the memory required to store 3D line information compared to the original 3D point clouds, making it more efficient. However, to obtain the 3D line cloud representation, the data must undergo a one-time permanent lifting transformation. This transformation is performed only once since performing it multiple times could potentially reveal the 3D points by intersecting them with the corresponding 3D lines. Additionally, while the recovery of a single 3D point from the 3D line representation presents a challenging inversion problem, there is a possibility to infer information about the scene structure through an analysis of the 3D line cloud's density. These results show that these specific challenges must be addressed to ensure its effectiveness in privacy protection. It needs to improve its accuracy and robustness to ensure that no traces of sensitive data can be reconstructed from the 3D line cloud representation. Despite its challenges, the main advantage of this approach is its scalability and computational benefits. As scenes become more complex or involve multiple cameras, 3D line cloud representations offer scalability and computational efficiency due to their reduced memory and processing, making it beneficial for frameworks like Hydra.

The conservative plane releasing for spatial privacy protection approach addresses the spatial inference threat. This framework combines spatial generalizations with conservative releasing to conceal the general location and the exact location of an agent by limiting the information that is shared and providing only sufficient spatial information to offer a generic representation of the spaces. To test this approach, developers designed an attacker that has previous knowledge of the environment and that aims to infer the inter-space and intra-space location of a user. For this purpose, they define privacy metrics to measure the accuracy of the estimations. The inter-space location is measured using misclassification error rates, and the intra-space location is measured with distance errors, which evaluate the accuracy in estimating a user's position within a space in meters. The results show that adversaries in scenarios with a radius of 0.5 and a maximum release of 17 planes will misidentify the inter-space location at least half of the time. Additionally, the estimation of the intra-space location will have a distance error of 3.0 meters. Similarly, they obtain similar results for scenarios with a radius of 1.0 and a maximum release of 11. These results show that the Quality-of-Service (QoS) varies with different factors. The increasing size of the radius has a small positive effect on the overall QoS, but it is more influenced by the number of planes. Smaller radius and fewer successive releases result in better QoS. Thus, in this approach there is a trade-off between prioritizing privacy and utility. This trade-off involves deciding how much information is shared while considering the QoS. Prioritizing privacy suggest limiting the information that is released, which reduces the chances of privacy breaches but may result in lower utility. On the other hand, prioritizing utility suggests revealing more information, enhancing data quality. To enhance privacy, one can use a larger radius, limiting the number of planes released and to improve utility, a smaller radius can be used, allowing more freedom in releasing a higher number of planes. Furthermore, this works suggests that plane generalizations that closely match the actual surfaces might maintain the user's experiences in AR applications. In summary, this approach effectively reduces the risk of revealing details of the environment; however, it comes with drawbacks such as a loss in accuracy and a trade-off in utility. The application of this approach into the Hydra framework could be studied while carefully considering the loss of accuracy in the representation.

5.4 Output protection

Lastly, after the processing stage, an output is generated to be rendered and exploited by applications. At this stage, users have no control over how their data is being used and for what purposes by these applications. Hydra dynamically reconstructs a comprehensive 3D representation of real-time environments, which can be leveraged in smart building scenarios. In line with the input stage, sensitive information may be captured in the process of building the model of the environment and may be present in the rendered output, which results in an intersection between input and output protection. In addition, untrusted third parties may gain access to the rendered outputs, which poses a risk of potential unauthorized modifications that could compromise the reliability of the outputs and give raise to privacy concerns [57][58].

5.4.1 Elicit mitigation strategies

This category can be associated with both branches from the mitigation techniques discussed in Section 3.2.2. Aligned with the input protection approaches, protection approaches in the output data stream can involve the removal of sensitive information. This approach focusses on guarding the exposure of sensitive data after it has been captured and rendered, ensuring that only necessary information is shared with the applications, which can be categorized as rendering approaches. These measurements can be employed as an intermediate layer, situated between the rendering and the applications interfaces. Other protection approaches focus on minimizing the exposure after the potential disclosure of sensitive information, which are categorized as output reliability and are also implemented as an intermediate layer [57][58].

5.4.2 Selection of PETs

Following the same methodology as for input and data protection, various approaches can be employed to ensure that data remains protected when it is presented to the users. These different approaches to output protection come with their own advantages and mainly focus on ensuring output reliability. Some of these approaches control output access with an object-level granularity [74] that manages the output rendering. Arya [75] is a following approach that establishes an output policy specification and enforcement, which was later improved by an approach that integrated reinforcement learning for dynamic environments [76].

ARYA

The Arya framework is one of the first approaches that explores the integration of an output policy module, aiming to solve some of the risks arising from malicious actors tampering with the output stream. Potential risks include the manipulation of the visual representation, such as the intentional obscuring of real-world content or the inclusion of visual content and misleading information. The main contribution of this approach is to address these risks that can potentially result in undesirable outputs, posing a security risk that can further lead to privacy implications. For this purpose, it aims to propose a policy specification framework to define output policies. These policies specify what actions applications are allowed to perform, dictating how they can access and interact with real-world objects or elements in a user's environment [75].

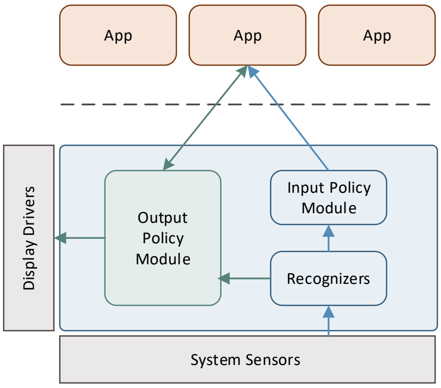


Figure 5.8: Arya's architecture [75].

This framework leverages on previous work of the system recognizers [64] and an input policy module [77] as shown in Figure 5.8. The recognizers module is used for gathering and interpreting raw sensor data from the real world and it exposes only the higher-level objects. Originally designed for input protection, Recognizers is able to identify specific elements within the raw sensor data like people, faces or flat surfaces. It promotes a least-privilege approach, ensuring that application access only the required information. In the context of Arya, it now offers valuable information about the user's physical environment to the output control policies [64].

In the first branch of the architecture, the Recognizers module is followed by an input module that is used to determine the selection of data to be passed to applications. This module automatically detects and enforce policies without requiring user involvement. In this approach, they define policies through "passports" that are presented by the objects, places, and individuals. These passports specify its concrete policies, being the system itself the one deciding which level of permission they are granting applications. Thus, this method simplifies the process of protecting information for users by initially defining policies for them based on context, yet it also grants them the flexibility to modify these policies as required. In some examples scenarios, like a workplace, it may be specified that whiteboards need protection, while locker rooms enforce a "no recording" policy [77].

In the second branch of the architecture, Arya integrates an output policy module before the output display that controls the visual output by aligning it with the specified policies. This design involves defining the desired output policies, implementing them effectively, managing situations where policies may clash and formulating countermeasures to address violations of those policies. Firstly, Arya aims to transform abstract guidelines into concrete polices by describing specific conditional predicate or Boolean expressions and mechanisms or actions. Each conditional predicate describes an undesired condition in the output display that determines when a policy should be implemented and then, a mechanism can be implemented to resolve it. Thus, policies are formed as a combination of one condition and one mechanism from the set of options provided by Arya instead of arbitrary policies. These policies can include avoiding the obscuring of real elements, abrupt movements of objects or objects that hide another object. Secondly, these policies are applied at different points in the Arya pipeline, during the creation or modification of objects or by monitoring policy conditions on every frame. Lastly, instances where policy violations are detected, are identified and partially shared with applications to address those violations. Additionally, Arya showcases its ability to effectively support policies from diverse sources while also considering the existence of malicious policies. In the case of conflicting policies, the less intrusive policy will prevail. This can result in applications displaying less content, potentially impacting in their functionality. However, they cannot result in a more intrusive output [75].

Adaptive Fog-Based Output Security for Augmented Reality

The adaptative fog-based output security framework builds upon Arya and enhances it by generating adaptative policies instead of them being bounded to a set of policies. The main contribution of this framework lies in its ability to automatically generate policies that adjust to changing conditions, making it suitable to dynamic environments. For this purpose, this framework leverages on deep reinforcement learning (RL) and fog computing, as shown in Figure 5.9 [76].

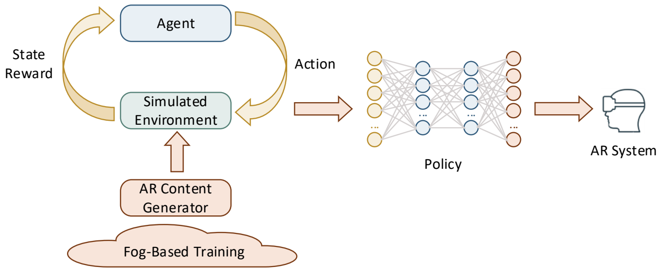


Figure 5.9: Pipeline for generating and deploying an output security policy for AR systems using deep reinforcement learning.[76].

In a general sense, deep reinforcement learning (RL) facilitates task automation. In this approach, deep RL is used for automating decision-making through interactions with a dynamic environment. As shown in Figure 5.9, the agent observes and interacts with the simulated environment. Based on its observations, the agent selects a policy to implement and in response it receives a reward function from the environment, which represents the effectiveness of the chosen policy. This measure is represented through an obstruction metric, which indicates the percentage of the user's display that is obstructed. Through this iterative process, the agent optimizes its decision-making and adjusts the neural network for implementing the optimal policy. Thus, this framework develops a "smart middlebox" system designed to filter out potentially harmful or unintended elements from the output stream before displaying them. This is achieved by performing offline training simulations through simulated environments to obtain a neural network model. Then, this model can be applied to real objects by making use of the system Recognizers. Some of the key benefits of RL is that it relies on environmental data and does not require generating new training sets, making it adaptable to different scenarios. Policies can be adjusted by modifying the reward function, allowing the agent to adapt its decisions. Additionally, because of its closed-loop nature, the model can consider long-term consequences of an agent's actions and can predict interference due to their movements [76].

Complementary to deep RL, fog computing brings computation closer to the end devices to optimize data processing. This localized processing reduces network latency, improves response times, and reduces dependency on centralized cloud servers. As shown in the architecture, fog computing is used for the training of the deep neural network. This design allows gathering more contextual information about the environment, enhancing accuracy and the realism of training simulations [76].

This framework has been tested through some experiments that show its effectiveness. Firstly, this system trains the agent in regulating the placing of holograms in AR scenarios, which significantly reduce obstruction of real-world objects in real-time scenarios. The obstruction metric was tested, showing that the average obstruction initially higher at around 34%, without policies. That value is significantly reduced to around 7% by applying policies that combine the displacement of objects and the incorporation of transparency policies. However, this metric alone cannot fully assess the effectiveness of AR content preservation, as policies can achieve an obstruction value of 0% by making objects transparent or deleting them. Thus, there is a need to develop alternative metrics. The reward function and the agent's value estimate are both tested to illustrate the improvement in the agent's performance over time. As the agent undergoes training, it becomes more skilled at predicting and anticipating future reward functions, reflecting its increasing skill level. Also, these experiments show that the RL generated policies had a minimal impact on frame rates. It shows that while activating the policy, there is a slight decrease in the frame rate compared to the non-policy case, while both performances converge as the number of holograms increases. This suggest that RL policies can be integrated into AR systems without compromising performance [76].

5.4.3 Output protection applied to Hydra

During the output stage, most of the potential threats from perceptual applications are centered around security concerns, which have the potential to compromise individual privacy. External parties that gain access to the output stream can manipulate the scene graphs by adding or removing visual content, thereby altering, or removing necessary information and introducing distraction for individuals. The threat identified in Table 4.3.3 as Threat 9 refer to external entities gaining access to the output stream. Ensuring the protection of this information is necessary to prevent additional threats and it can be achieved by implementing the previously discussed frameworks.

Arya proposes an approach that addresses both input protection, achieved by integrating Recognizers and the input policy module, and output protection, achieved by integrating Recognizers with the output policy module. In this section, I will focus on the branch dedicated to output protection. Here, Recognizers identifies real-world objects and their positions, enabling the output policy module to enforce policies based on this information. Arya mitigates the risks associated with obscuring real-world content or introducing unintended visual content into the output stream. It facilitates the process of defining and enforcing output policies, as well as resolving conflicts among these policies. For example, if Recognizers detect a person nearby, the module can ensure that AR objects do not occlude or obstructs that person's view. Thus, Arya focuses on enforcing policies at a level of individual objects, without affecting the entire output and supporting operations like movement and resizing. One of the drawbacks of this approach is that it relies on a set of predefined policies, which poses a significant challenge for frameworks like Hydra that operate in dynamic and unpredicted environments. Hydra requires dynamic, real-time generation of visual content and generic policies may fall short in addressing the finer, context-specific requirements, potentially limiting its adaptability and flexibility. Furthermore, the defined policies in this approach focus on AR applications and are customized to handle scenarios specific to the AR context. To successfully integrate them into the Hydra pipeline, they will have be to be adapted. Some of these policies, such as "Avoid abrupt movement of AR objects" or "Don't display text messages or social media while driving" are centered around AR scenarios. Other policies, like "Don't obscure pedestrians or road signs" focus on outdoor scenarios that cannot be directly applied to Hydra. However, some policies like "Don't obscure exit signs" could potentially be adapted to Hydra's scenarios. Moreover, more suitable policies should be defined to adapt to Hydra's specific context and requirements, ensuring that they align with the unique challenges and objectives of the framework. This involves tailoring policies to address dynamic indoor environments, diverse sensor inputs, and the need for real-time content generation and adaptation. Additionally, this approach depends on Recognizers for accurate detection. Failures or inaccuracies in detection can result in incorrect policy implementation. These inaccuracies may occur due to various factors, including noisy sensor input or challenging real-world conditions.

Attempting to create a comprehensive set of rule-based policies that anticipates and covers every possible situation in the complexity of dynamic scenarios is extremely challenging. For this reason, the adaptative fog-based output security builds upon Arya and proposes the automatic generation of complex policies. This framework allows policies to adapt and respond to dynamic environments with the purpose of protecting the output stream. Developers conducted tests on this framework to show that the agent's performance improves over time, enabling it to effectively regulate hologram placement and significantly reduce the obstruction of real objects. However, this approach primarily measures the correct placement of holograms by using an obstruction metric. This metric represents the percentage of obstruction on the user's display, but it can be misleading since a low percentage may result from either directly removing information or making it transparent. Therefore, relying solely on this metric may not fully assess its effectiveness in AR scenarios that aim to integrate external information in the users view. On the contrary, Hydra aims to serve as a true representation of the environment, shifting its focus to detecting virtual content and removing what is not real and accurate within the environment. In this scenario, the requirements are simplified, focusing in removing all content external to reality. Consequently, the obstruction metric for Hydra's functionality may quantify the percentage of elements that are external to reality and need to be removed, serving as a metric for assessing the realism of the representation. Additionally, tests show that enabling these policies reduces the frame rate by only 10-12 frames per second (fps), resulting in an operational frame rate of nearly 80 fps. This reduction does not have a great impact on real-time representations, especially when compared to the standard frame rate of 24 fps in media. In summary, these advancements work towards achieving a possible integration into real-time scenario, given its adaptability to dynamic environments, flexibility in policy adjustment and long-term planning capabilities. This suitably extends to scenarios where Hydra may be integrated.

6. Conclusions and future directions

6.1 Conclusions

In this work, I have delved into the privacy implications of integrating Hydra's novel framework into smart building domains. The exploration began with a literature review encompassing key subjects: privacy, computer vision, and smart buildings. Subsequently, the main focus shifted towards privacy threat modeling, where I chose the LINDDUN methodology due to its primary focus on privacy concerns, aligning with the main focus in this thesis. While this method was initially designed for web applications, I have explored its applicability in newer technologies like 3DSG. Following the two main sets of steps in this approach, I have analyzed potential threats that may arise and identified potential PETs that could mitigate them. During the course of this study, I have achieved significant outcomes. Firstly, I have developed two DFDs that provide insights into Hydra's functionality at various levels of abstraction, represented by the level 0 DFD and the level 1 DFD. Secondly, I have compiled a comprehensive set of potential privacy threats. These threats are organized into a table, aligning each threat with the corresponding LINDDUN properties and the DFD elements that may be susceptible to these threats. Lastly, I have proposed several PETs as mitigation technologies to address these threats. The analysis explores the opportunities and challenges associated with the potential integration of these PETs into the Hydra framework. Collectively, these outcomes contribute to a deeper understanding of the potential privacy risks and mitigation strategies within the integration of Hydra's framework into the context of smart buildings. In summary, throughout this study, I have emphasized the significance of conducting privacy threat analysis for emerging technologies prior to their implementation. Furthermore, I have highlighted the necessity to adapt and develop PETs tailored to address potential threats in these scenarios.

6.2 Future directions

The ultimate goal of Hydra is to construct an accurate representation of the environment, aiming to emulate human perception for extracting valuable information. However, this framework is still under development and requires testing in practical scenarios. Throughout this project, I have emphasized the importance of privacy, particularly within the field of 3DSG. In this field, the predominant focus has been centered on advancing the capabilities and potential of 3DSG technology. It is important to highlight that there has been limited research addressing the associated privacy implications.

This line of research can be further extended, as these novel representations could benefit from the data collected by various sensors integrated into a smart building. Information from temperature, air quality, and more can potentially contribute to these representations to enrich its functionality and to improve the automation of processes. However, with this expansion of capabilities privacy concerns naturally escalate. Consider, for instance, the integration of room temperature data to enhance the representation and automation within these environments. A rise in temperature could be indicative of human presence in a room. Such privacy related aspects could be further explored in future research to protect user privacy.

Bibliography

[1] Nathan Hughes, Yun Chang, and Luca Carlone. "Hydra: A Real-time Spatial Perception System for 3D Scene Graph Construction and Optimization". In: *Robotics: Science and Systems*. Massachusetts Institute of Technology. New York City, NY, USA, June 2022.

[2] *What is GDPR, the EU's new data protection law?* URL: `https://gdpr.eu/what-is-gdpr/`.

[3] *International Organization for Standardization (ISO) Standards*. URL: `https://www.iso.org/standards.html`.

[4] Edna Dias Canedo et al. "Privacy requirements elicitation: a systematic literature review and perception analysis of IT practitioners". In: *Requirements Engineering* 28 (2023), pp. 177–194. DOI: `10.1007/s00766-022-00382-8`. URL: `https://link.springer.com/article/10.1007/s00766-022-00382-8`.

[5] International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC). *Information technology — Security techniques — Privacy engineering for system life cycle processes*. Technical Report ISO/IEC TR 27550:2019. ISO/IEC, 2019. URL: `https://standards.iteh.ai/catalog/standards/sist/f90e3679-afab-4fa9-971d-7317f92ed8c1/iso-iec-tr-27550-2019`.

[6] National Institute of Standards and Technology (NIST). *LINDDUN Privacy Threat Modeling Framework*. 2023. URL: `https://www.nist.gov/privacy-framework/linddun-privacy-threat-modeling-framework`.

[7] KU Leuven DistriNet Research Unit. *LINDDUN: Privacy Threat Modeling*. 2021. URL: `https://linddun.org/` (visited on 2023).

[8] Jeroen van den Hoven et al. *Privacy and Information Technology*. First published Thu Nov 20, 2014; substantive revision Wed Oct 30, 2019. Stanford Encyclopedia of Philosophy. Oct. 2019. URL: `https://plato.stanford.edu/entries/it-privacy/#PerDat`.

[9] *NIST Computer Security Resource Center (CSRC) Glossary*. National Institute of Standards and Technology (NIST). URL: `https://csrc.nist.gov/glossary/term/PII`.

[10] *GDPR Personal Data*. URL: `https://gdpr-info.eu/issues/personal-data/#:~:text=GDPR%20Personal%20Data&text=4%20(1).,identified%20or%20identifiable%20natural%20person`.

[11] *Art. 6 GDPR Lawfulness of processing*. URL: `https://gdpr.eu/article-6-how-to-process-personal-data-legally/`.

[12] Ann Cavoukian. *Privacy by Design*. Information & Privacy Commissioner, Ontario, Canada.

[13] *Art. 25 GDPR: Data protection by design and by default*. URL: `https://gdpr.eu/article-25-data-protection-by-design/`.

[14] *Art. 7 GDPR: Conditions for consent*. URL: `https://gdpr.eu/article-7-how-to-get-consent-to-collect-personal-data/`.

[15] *Art. 5 GDPR: Principles relating to processing of personal data.* URL: `https://gdpr.eu/article-5-how-to-process-personal-data/`.

[16] International Organization for Standardization. *Information technology — Security techniques — Information security management systems — Overview and vocabulary.* International Standard ISO/IEC 27000:2018(E). ISO, Feb. 2018.

[17] International Organization for Standardization. *ISO/IEC 27000 family.* Accessed on Month Day, Year. 2023. URL: `https://www.iso.org/standard/iso-iec-27000-family`.

[18] International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC). *Information technology — Security techniques — Privacy framework.* First edition, 2011-12-15. 2011.

[19] IBM. *What is computer vision?* Accessed on Month Day, Year. 2023. URL: `https://www.ibm.com/topics/computer-vision`.

[20] Richard Szeliski. *Computer Vision: Algorithms and Applications.* The University of Washington, 2022.

[21] Helen Oleynikova et al. "Signed Distance Fields: A Natural Representation for Both Mapping and Planning". In: (2017). Autonomous Systems Lab, ETH Zurich.

[22] Petter Bogen Sydhagen. "How Can We Distinguish Perception from Cognition? The Perceptual Adaptation Hypothesis". Master's Thesis. University of Oslo, 2017.

[23] David J. Chalmers, Robert M. French, and Douglas R. Hofstadter. *High-Level Perception, Representation, and Analogy: A Critique of Artificial Intelligence Methodology.* Tech. rep. CRCC Technical Report 49. Bloomington, Indiana 47408: Center for Research on Concepts and Cognition, Indiana University, Mar. 1991.

[24] Guangming Zhu et al. "Scene Graph Generation: A Comprehensive Survey". In: (June 2022). arXiv: `arXiv:2201.00443v2 [cs.CV]`.

[25] Antoni Rosinol et al. "Kimera: from SLAM to Spatial Perception with 3D Dynamic Scene Graphs". In: *The International Journal of Robotics Research* (2021).

[26] Antoni Rosinol et al. "Kimera: from SLAM to Spatial Perception with 3D Dynamic Scene Graphs". In: *The International Journal of Robotics Research* (2021). Accepted for publication at IJRR 2021.

[27] Helen Oleynikova et al. "Sparse 3D Topological Graphs for Micro-Aerial Vehicle Planning". In: *2016 IEEE International Conference on Robotics and Automation (ICRA).* IEEE. 2016.

[28] A.H. Buckman, M. Mayfield, and Stephen B.M. Beck. "What is a Smart Building?" In: *Emerald* (2012). URL: `www.emeraldinsight.com/2046-6099.htm`.

[29] David Eckhoff and Isabel Wagner. "Privacy in the Smart City—Applications, Technologies, Challenges, and Solutions". In: *IEEE Communications Surveys & Tutorials* 20.1 (2018), p. 489. DOI: `10.1109/COMST.2017.2769384`.

[30] Jie Lin et al. "A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications". In: *IEEE Internet of Things Journal* 4.5 (Oct. 2017), p. 1125.

[31] Tomás Domínguez-Bolaño et al. "An overview of IoT architectures, technologies, and existing open-source projects". In: *Internet of Things* 20 (Oct. 2022), p. 100626. ISSN: 2542-6605. DOI: `10.1016/j.iot.2022.100626`. URL: `https://www.sciencedirect.com/science/article/pii/S2542660522001271`.

[32] Michael Cobb. *Threat Modeling.* 2021. URL: `https://www.techtarget.com/searchsecurity/definition/threat-modeling` (visited on 2023).

[33] Nicolas Montauban. *What is threat modeling? A conversation with two experts.* 2023. URL: `https://codific.com/privacy-threat-modeling/#:~:text=Privacy%5C%20threat%5C%20modeling%5C%20is%5C%20a,threats%5C%20and%5C%20protect%5C%20data%5C%20privacy` (visited on 2023).

[34] Nataliya Shevchenko et al. "Threat modeling: A summary of available methods." In: (July 2018).

[35] *STRIDE Threat Modeling: What You Need to Know.* URL: `https://www.softwaresecured.com/stride-threat-modeling/`.

[36] threatmodeler. *Threat modeling methodologies: What is VAST?* Oct. 2018. URL: `https://threatmodeler.com/threat-modeling-methodologies-vast/`.

[37] Kim Wuyts and Wouter Joosen. *LINDDUN privacy threat modeling: a tutorial.* Technical Report. CW Reports. Department of Computer Science, KU Leuven, July 2015.

[38] KU Leuven DistriNet Research Unit. *LINDDUN Pro Privacy Threat Modeling Tutorial.* 2023. URL: `https://downloads.linddun.org/tutorials/pro/v0/tutorial.pdf` (visited on 2023).

[39] KU Leuven DistriNet Research Unit. *Privacy Threat Types.* 2023. URL: `https://linddun.org/threat-types/` (visited on 2023).

[40] Andreas Pfitzmann and Marit Hansen. *Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology.* `https://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.31.pdf`. PDF file. 2008.

[41] University of Florida. *Creating an Information System/Data Flow Diagram.* URL: `https://security.ufl.edu/resources/risk-assessment/creating-an-information-systemdata-flow-diagram/` (visited on 2023).

[42] SBS Cybersecurity. *DATA FLOW DIAGRAMS 101.* 2023. URL: `https://sbscyber.com/resources/data-flow-diagrams-101` (visited on 2023).

[43] *Levels in Data Flow Diagrams (DFD).* 2023. URL: `https://www.geeksforgeeks.org/levels-in-data-flow-diagrams-dfd/` (visited on 2023).

[44] Michael Howard and Steve Lipner. *The Security Development Lifecycle.* Microsoft Press, 2006.

[45] OWASP. *Risk Asessment Framework.* URL: `https://owasp.org/www-project-risk-assessment-framework/` (visited on 2023).

[46] Microsoft. *Improving web application security: Threats and countermeasures.* 2010. URL: `https://learn.microsoft.com/en-us/previous-versions/msp-n-p/ff648644(v=pandp.10)?redirectedfrom=MSDN#c03618429_011` (visited on 2023).

[47] NIST. *Guide for Conducting Risk Assessments.* 2012. URL: `https://csrc.nist.gov/publications/detail/sp/800-30/rev-1/final` (visited on 2023).

[48] Software Engineering Institute. *Introducing OCTAVE Allegro: Improving the Information Security Risk Assessment Process.* 2007. URL: `https://resources.sei.cmu.edu/library/asset-view.cfm?assetID=8419` (visited on 2023).

[49] E. Zio. "The future of risk assessment". In: *Reliability Engineering & System Safety* 177 (Sept. 2018), pp. 176–190. DOI: `10.1016/j.ress.2018.04.025`. URL: `https://www.sciencedirect.com/science/article/pii/S0951832017301359`.

[50] OWASP (Open Web Application Security Project). *OWASP Risk Rating Methodology.* 2023. URL: https://owasp.org/www-community/OWASP_Risk_Rating_Methodology.

[51] Martina Brachmann et al. “Toward Privacy-Preserving Localization and Mapping in eXtended Reality: A Privacy Threat Model”. In: *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit).* 2023, pp. 635–640. DOI: 10.1109/EuCNC/6GSummit58263.2023.10188227.

[52] Jie Lin et al. “A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications”. In: *IEEE Internet of Things Journal* 4.5 (Oct. 2017), pp. 1125–1142. DOI: 10.1109/JIOT.2017.2763839.

[53] Sunho Kim et al. “Extending Data Quality Management for Smart Connected Product Operations”. In: *IEEE Access* 7 (2019), pp. 138598–138616. DOI: 10.1109/ACCESS.2019.2945124.

[54] Brittan Heller. “Watching Androids Dream of Electric Sheep: Immersive Technology, Biometric Psychography, and the Law”. In: *Vanderbilt Journal of Entertainment and Technology Law* 23 (1 2021). URL: https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss1/1.

[55] Francesco Pittaluga et al. “Revealing Scenes by Inverting Structure from Motion Reconstructions”. In: *arXiv preprint arXiv:1904.03303* (Apr. 2019).

[56] Alessandro Acquisti, Ralph Gross, and Fred Stutzman. “Face Recognition and Privacy in the Age of Augmented Reality”. In: *Journal of Privacy and Confidentiality* 6.2 (2014), pp. 1–20.

[57] Jaybie A. De Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. “Security and Privacy Approaches in Mixed Reality: A Literature Survey”. In: *arXiv preprint arXiv:1802.05797* (June 2020). arXiv: 1802.05797 [cs.CR].

[58] Jaybie A. De Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. “Privacy and Security Issues and Solutions for Mixed Reality Applications”. In: *Springer Handbook of Augmented Reality.* Springer, 2023, pp. 123–145. DOI: 10.1007/978-3-030-67822-7. URL: https://link.springer.com/book/10.1007/978-3-030-67822-7.

[59] Suman Jana, Arvind Narayanan, and Vitaly Shmatikov. “A Scanner Darkly: Protecting User Privacy From Perceptual Applications”. In: *2013 IEEE Symposium on Security and Privacy.* IEEE, 2013, pp. 349–363. DOI: 10.1109/SP.2013.31.

[60] Eisa Zarepour et al. “A Context-based Privacy Preserving Framework for Wearable Visual Lifeloggers”. In: *2016 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops).* IEEE. 2016.

[61] Nisarg Raval et al. “MarkIt: privacy markers for protecting visual secrets”. In: *UbiComp ’14 Adjunct: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication.* ACM. 2014, pp. 1289–1295. DOI: 10.1145/2638728.2641707.

[62] Jiayu Shu, Rui Zheng, and Pan Hui. “Cardea: Context-Aware Visual Privacy Protection from Pervasive Cameras”. In: *arXiv preprint arXiv:1610.00889* (2016). arXiv: 1610.00889 [cs.CR].

[63] Aakash Shrestha et al. “Virtual Curtain: A Communicative Fine-grained Privacy Control Framework for Augmented Reality”. In: *2023 International Conference on Computing, Networking and Communications (ICNC): Communications and Information Security Symposium.* IEEE. 2023.

[64] Suman Jana et al. "Enabling Fine-Grained Permissions for Augmented Reality Applications with Recognizers". In: *Proceedings of the 22nd USENIX Security Symposium.* Washington, D.C., USA: USENIX Association, Aug. 2013. ISBN: 978-1-931971-03-4.

[65] Lucas Silva Figueiredo et al. "Prepose: Privacy, Security, and Reliability for Gesture-Based Programming". In: May 2016, pp. 122–137. DOI: 10.1109/SP.2016.16.

[66] Jaybie Agullo de Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. "SafeMR: Privacy-aware Visual Information Protection for Mobile Mixed Reality". In: *2019 IEEE 44th Conference on Local Computer Networks (LCN).* 2019, pp. 254–257. DOI: 10.1109/LCN44214.2019.8990850.

[67] Open Source Computer Vision. *OpenCV Documentation - Introduction to OpenCV.* 2023. URL: https://csrc.nist.gov/publications/detail/sp/800-30/rev-1/final (visited on 2023).

[68] Chao-Yong Hsu and Chun-Shien Lu. "Homomorphic Encryption-based Secure SIFT for Privacy-Preserving Feature Extraction". In: *Proceedings of SPIE - The International Society for Optical Engineering* 7880 (Feb. 2011). DOI: 10.1117/12.873325.

[69] Linzhi Jiang et al. "Secure outsourcing SIFT: Efficient and Privacy-Preserving Image Feature Extraction in the Encrypted Domain". In: *IEEE Transactions on Dependable and Secure Computing* PP (Sept. 2017), pp. 1–1. DOI: 10.1109/TDSC.2017.2751476.

[70] Zhan Qin et al. "SecSIFT: Secure image SIFT feature extraction in cloud computing". In: *ACM Transactions on Multimedia Computing, Communications, and Applications* 12 (Sept. 2016). DOI: 10.1145/2978574.

[71] Pablo Speciale et al. "Privacy Preserving Image-Based Localization". In: *arXiv preprint arXiv:1903.05572v1* (Mar. 2019). arXiv: 1903.05572v1 [cs.CV].

[72] Jaybie Agullo de Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. "Conservative Plane Releasing for Spatial Privacy Protection in Mixed Reality". In: *arXiv preprint arXiv:2004.08029* (2020). arXiv: 2004.08029 [cs.CV].

[73] Jaybie A. de Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. "A First Look into Privacy Leakage in 3D Mixed Reality Data". In: *European Symposium on Research in Computer Security.* Springer, 2019, pp. 149–169.

[74] Kiron Lebeck, Tadayoshi Kohno, and Franziska Roesner. "How to Safely Augment Reality: Challenges and Directions". In: *Proceedings of the 17th International Workshop on Mobile Computing Systems and Applications.* Feb. 2016, pp. 45–50. DOI: 10.1145/2873587.2873595.

[75] Kiron Lebeck et al. "Securing Augmented Reality Output". In: *Proceedings of the 38th IEEE Symposium on Security and Privacy (Oakland).* 2017. URL: https://ar-sec.cs.washington.edu.

[76] Surin Ahn et al. "Adaptive Fog-Based Output Security for Augmented Reality". In: *Proceedings of the 2018 Morning Workshop on Virtual Reality and Augmented Reality Network (VR/AR Network '18).* 2018, pp. 1–6. DOI: 10.1145/3229625.3229626. URL: https://doi.org/10.1145/3229625.3229626.

[77] Franziska Roesner et al. "World-Driven Access Control for Continuous Sensing". In: *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14).* Nov. 2014, pp. 1169–1181. DOI: 10.1145/2660267.2660319. URL: https://doi.org/10.1145/2660267.2660319.

[78] Antoni Rosinol et al. “Kimera: from SLAM to Spatial Perception with 3D Dynamic Scene Graphs”. In: *The International Journal of Robotics Research* (2021). Accepted for publication at IJRR 2021.

[79] OWASP. *Risk Rating Methodology.* URL: `https://owasp.org/www-community/OWASP_Risk_Rating_Methodology` (visited on 2023).

A. Annex A: References to the LINDDUN threat trees.

A.1 LINDDUN threat trees level-0 DFD.

Table A.1: Linkability threat tree level-0 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	E2	DF1	L.2.1.1	Quasi-identifier combining data of a single individual. Location trace.	1
E2	DF2	E1	DF2	L.1.1	Unique identifier. User IDs.	2
E2	DF2	E1	DF2	L.2.2.1	Profiling an individual. Analyzing timing patterns.	1
E1	DF3	E3	DF3	L.2.2.2	Profiling a group of individuals. Energy consumption meter.	3
P2	DF6	E1	DF6	L.2.2.1	Profiling an individual. Requests to external services and analysing timing patterns.	4
P1	DF7	P2	DF7	L.2.2.1	Profiling an individual. Requests to external services and analysing timing patterns.	5
P2	DF8	P1	DF8	L.2.2.1	Profiling an individual. Requests to external services and analysing timing patterns.	5
P2	DF9	DS3	P2, DS3	L.2.2.2	Profiling a group of individuals. Process and storage.	6

Table A.2: Identifiability threat tree level-0 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	E2	E1, DS2	I.2.2	Revealing attributes.	7
E2	DF2	E1	DF2	I.2.1.1	Unique identifier. User IDs.	2
E1	DF5	P2	E1, DF5	I.1.2	Identified information in metadata.	7
P2	DF9	DS3	P2, DS3	I.2.2	Revealing attributes	7

Table A.3: Non-repudiation threat tree level-0 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	E2	DF1	Nr. 1.1	Attributable data evidence. Logged transmissions.	1
E2	DF2	E1	DF2	Nr.1.1	Attributable data evidence. Logged transmissions.	1
P2	DF6	E1	DF6	Nr.1.1	Attributable data evidence. Logged transmissions.	4
P1	DF7	P2	DF7	Nr.1.1	Attributable data evidence. Logged transmissions.	5
P2	DF8	P1	DF8	Nr.1.1	Attributable data evidence. Logged transmissions.	5
P2	DF9	DS3	P1, P2	Nr.2	Attributable action side-effect evidence. Action logging.	8

Table A.4: Detectability threat tree level-0 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	E2	DF1	D.1	Observed communications.	1
E2	DF2	E1	DF2	D.1	Observed communications.	1
P2	DF6	E1	DF6	D.1	Observed communications.	4
P1	DF7	P2	DF7	D.1	Observed communications.	5
P2	DF8	P1	DF8	D.1	Observed communications.	5
P1	DF7	P2	P1, P2	D.2	Application side-effect.	8

Table A.5: Disclosure of information threat tree level-0 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	E2	E1, DS2	DD.1.2	Disclosed personal data is fine-grained level of granularity.	7
E1	DF3	E3	DF3	DD.4.1.1	Predetermined set of parties. Use of third-party tracking and analytics services.	3
P1	DF7	P2	DF7	DD.3.2	Propagation of sensitive information.	5
P2	DF8	P1	DF8	DD.3.2	Propagation of sensitive information.	5
P2	DF9	DS3	P2, DS3	DD.1.2	Disclosed personal data is fine-grained level of granularity.	7

Table A.6: Unawareness threat tree level-0 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	E2	E1, DF1	U.2.1	Lack of data subject control. Preferences.	9
E1	DF3	E3	DF3, E3	U.1.1	Unawareness as data subject. Third parties.	10
P1	DF5	P2	DF5	U.2.1	Lack of data subject control. Preferences.	9
P2	DF6	E1	DF6	U.2.1	Lack of data subject control. Preferences.	9
P1	DF7	P2	P1, DS1, DF7	U.2.1	Lack of data subject control. Preferences.	11
P2	DF8	P1	DF8	U.2.1	Lack of data subject control. Preferences.	11
P2	DF9	DS3	P2, DF9, DS3	U.2.1	Lack of data subject control. Preferences.	12

Table A.7: Non.compliance threat tree level-0 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	E2	E1, DF1	Nc.2	Improper personal data management.	13
E1	DF3	E3	DF3, E3	Nc.2	Improper personal data management.	13
P1	DF5	P2	DF5, P2	Nc.2	Improper personal data management.	13
P2	DF6	E1	DF6	Nc.2	Improper personal data management.	13
P1	DF7	P2	P1, DS1, DF7	Nc.2	Improper personal data management.	13
P2	DF8	P1	DF8	Nc.2	Improper personal data management.	13

A.2 LINDDUN threat trees level-1 DFD.

Table A.8: Linkability threat tree level-1 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	P1 P2 P6	DF1	L.2	Linkable data.	1
E1	DF1	P1 P2 P6	DF1	L.2.1.1	Quasi-identifier combining data of a single individual. Location trace.	2
E2	DF2	P6	DF2	L.2.1.1	Quasi-identifier combining data of a single individual. Location trace.	2
P1	DF3	DS1	DS1, DF3	L.2	Linkable data.	1
P1	DF3	DS1	DS1	L.2.1.1	Quasi-identifier combining data of a single individual. Location trace.	2
P4	DF8	DS3 P7	DS3, DF8	L.2.2.2	Profiling a group of individuals.	3
P5	DF11	DS5 P7	DS5, DF11	L.2.1	Linking through combination.	4
P7	DF13	DS6 P8	DS6, DF13	L.2.1	Linking through combination.	5
P8	D14	DS7	DS7, DF14	L.2.1	Linking through combination.	5

Table A.9: Identifiability threat tree level-1 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	P1 P2 P6	DF1	I.2.3	Data subject is distinguishable from others. Face recognition.	6
P1	DF3	DS1	DS1, DF3	I.2.3	Data subject is distinguishable from others. Face recognition.	6
P7	DF13	DS6 P8	DS6, DF13	I.2.1.2	Identifiable information. Quasi-identifier.	7
P8	D14	DS7	DS7, DF14	I.2.1.2	Identifiable information. Quasi-identifier.	7

Table A.10: Non-repudiation threat tree level-1 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	P1 P2 P6	DF1	Nr. 1.1	Attributable data evidence. Images.	6
E1	DF1	P1 P2 P6	DF1	Nr. 1.3	Attributable data evidence. Metadata.	2
E2	DF2	P6	DF2	Nr. 1.3	Attributable data evidence. Metadata.	2
P1	DF3	DS1	DS1, DF3	Nr. 1.1	Attributable data evidence. Images.	6
P1	DF3	DS1	DS1, DS3	Nr. 1.3	Attributable data evidence. Metadata.	2
P7	DF13	DS6 P8	DS6, DF13	Nr. 1.1	Attributable data evidence. Scene Graph.	5, 7
P8	DF14	DS7	DS7, DF14	Nr. 1.1	Attributable data evidence. Scene Graph.	5, 7

Table A.11: Detectability threat tree level-1 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	P1 P2 P6	DF1	D.1	Observed communications from external entities	8
E2	DF2	P6	DF2	D.1	Observed communications from external entities	8
E1	DF1	DS7	DF14	D.1	Observed communications from external entities	8

Table A.12: Disclosure of information threat tree level-1 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	P1 P2 P6	DF1	DD.2.1	Excessive amount of information collected	1
E1	DF1	P1 P2 P6	DF1	DD.3.2	Propagation of sensitive information	9
P8	DF14	DS7	DF14	DD.3.2	Propagation of sensitive information	9

Table A.13: Unawareness threat tree level-1 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
E1	DF1	P1 P2 P6	E1	U.2.1	Lack of subject control. Preferences.	10
E1	DF1	P1 P2 P6	E1	U.2.1	Lack of subject control. Preferences.	10

Table A.14: Non-compliance threat tree level-1 DFD.

S	DF	D	Location	Characteristics	Impact	Threat
			Whole system	Nc.2	Improper personal data management	11