Understanding the Concepts behind Ethics in Designing Cyber-Physical Systems

Yelyzaveta Kurkchi ✉, Ivan Compagnucci, and Catia Trubiani

Gran Sasso Science Institute, L'Aquila, Italy
{yelyzaveta.kurkchi,ivan.compagnucci,catia.trubiani}@gssi.it

Abstract. The widespread adoption of Cyber-Physical Systems (CPS) in real-world infrastructures raises significant challenges, e.g., in the healthcare domain, managing patient data is not trivial. Still, humans are not sufficiently involved in making decisions about their data, which poses a significant risk to their trust. Our research focuses on understanding the ethical concerns that should be treated as first-class citizens in the design of CPS. To this end, this paper proposes a novel conceptual approach, namely DEsign TEchniques for Cyber-Physical Ethical Systems (DeTe4CPES), revisiting traditional CPS specifications to incorporate ethics-based requirements as drivers for taking design decisions. A selection of ethics-based requirements (i.e., privacy, security, and transparency) is considered to extract the design objectives and techniques defined in recent literature. The novelty lies in modeling the identified concepts using knowledge graphs, thereby providing a conceptual view for software designers willing to address ethics-based concerns in CPS. A Business Process model is provided to illustrate a concrete use case for intertwining ethics-based requirements, and a prototype implementation enables the specification of evaluation metrics.

Keywords: Conceptual Modeling · Cyber-Physical Systems · Ethics · Software Design · Business Processes.

1 Introduction

Cyber-Physical Systems (CPS) emerge in the literature as an evolution of embedded systems, driven by a renewed interplay between software and hardware components, which triggers novel research challenges and opportunities [15]. CPS play a fundamental role in critical real-world infrastructures, such as production monitoring [13] and healthcare, where they support hospital services, as confirmed by a recent European Parliament study [10]. CPS need to respond to the evolving society, and it becomes necessary to involve humans in expressing their preferences. This focus is motivated by the nature of CPS, which often process sensitive data and trigger decisions that may directly affect humans.

Ethics of CPS represents the request raised by the European Parliament Scientific Foresight study [10], which identifies a set of key issues, such as the need to guarantee patients' data privacy. In the field of conceptual modeling,

grounded principles have been defined to capture and operationalize abstract domains, such as ethics, in which moral values must be faithfully embedded into system designs [23]. Building on this foundation, recent studies tackle ethical requirements and guide towards morally aware systems [9,12]. To the best of our knowledge, the proposed methodologies [9,12] provide modeling constructs targeting ethics-based requirements, however, they still lack the specification of the design techniques that implement those requirements.

This paper fosters the concept of *Cyber-Physical Ethical Systems* (CPES) in response to the European Parliament Scientific Foresight Study [10], in which *Privacy*, *Security*, and *Transparency* have been indicated as the primary ethical concerns in the design of CPS. Our goal is to introduce a conceptual approach, namely DEsign TEchniques for Cyber-Physical Ethical Systems (DeTe4CPES), that supports the design of CPS, so that these systems respect individuals' rights by promoting privacy safeguards, introducing proper security measures, and guaranteeing transparency in data collection and usage. To this end, we have conducted a literature review to identify weaknesses in current CPS design practices. Regarding conceptual modeling, we draw inspiration from the recent work of Rezayat et al. [32], which proposes the Hippocratic Application Programming Interfaces (APIs) as a solution to protecting data subjects' rights. Similar to [32], we adopt knowledge graphs [33] to model design techniques and their relationships, but our focus is on CPS. A concrete application of the approach is illustrated through a patient monitoring use case modeled using the well-established Business Process Model and Notation (BPMN) [26]. Business activities are associated with ethical shortcomings and mapped to design objectives and techniques, thereby fostering their quantitative evaluation.

Objectives of the paper. We consider privacy, security, and transparency as the primary requirements in designing CPS, and we extract from the literature the design objectives and techniques that contribute to the decision-making process. The main contributions of this paper are: (i) the identification of ethical shortcomings in current CPS design practices; (ii) the modeling of design objectives and techniques, structured as knowledge graphs, which contribute to the definition of DeTe4CPES; (iii) the application of DeTe4CPES to a concrete use case (i.e., patients equipped with wearable devices tracking their health status and triggering intervention of medical staff), in which a business process model shows how knowledge graphs can support the integration of ethics-based requirements into CPS design; and (iv) the prototype implementation of DeTe4CPES, i.e., a tool software designers can use to select techniques and specify evaluation metrics that quantify them. Replication data is available [18].

The paper is organized as follows. Section 2 positions our contribution with respect to related work. Section 3 briefly presents the background knowledge. Section 4 reports the main shortcomings for privacy, security, and transparency, and describes the research methodology. Section 5 presents the identified design objectives and techniques. Section 6 illustrates our conceptual approach through business processes, and Section 7 discusses the current limitations. Section 8 concludes the paper by outlining future research directions.

2 Related Work

We discuss related work starting with conceptual approaches, CPS design methodologies, and how current work embeds ethical principles into CPS design.

Conceptual approaches. A conceptual model of a pediatric disease is presented in [2], which is used to build a web application that assists clinicians in managing patients. Data warehousing has been analyzed using a conceptual model in [29] to capture analyzable data and analysis parameters, thereby evaluating the relationships among these parameters. An example is provided of a film company releasing movies, illustrating how the conceptual model supports the contract schema that cinema halls will analyze. A conceptual model for Internet of Things (IoT) applications is proposed in [30]. It gathers the constituent elements of such applications, i.e., the devices and the data they exchange, and uses them to monitor and regulate properties such as sensor temperature. A conceptual method is proposed in [1] to address trust concerns, as well as the connected workarounds, such as trustworthiness facets and goals. The objective is to elicit trust requirements and offer a user-centered software solution that respects users' concerns. The method is applied to a dating application, underscoring the need to verify the authenticity of user profiles. A conceptual method for modeling security requirements is proposed in [28]. It provides a high-level representation of assets and models actors' goals, information, and documents. DeTe4CPES builds a design-oriented knowledge graph that links selected ethics-based concerns to reusable design techniques that complement modeling techniques [28].

Cyber-Physical Systems. CPS design strategies must model both physical and digital aspects to effectively manage these systems [20]. The importance of conceptual models for structuring requirements engineering and guiding CPS design is discussed in [36]. A five-layer CPS design architecture for Industry 4.0 manufacturing, including data collection, processing, and analytics, is presented in [21]. A security-focused framework that incorporates cyber-attack failures into safety analysis is proposed in [34]. CPS-specific assets, such as attack trees, are analyzed in [31] to guarantee security at the requirements phase. A model-driven approach that uses Petri nets to detect errors early and enhance CPS reliability is presented in [14]. Our attempt to pursue Cyber-Physical Ethical Systems (CPES) is motivated by studies that highlight ethics as a novel concern in CPS design [10,22]. A literature review including 68 studies [22] found that papers on ethics in industrial CPS have nearly doubled, highlighting privacy, security, and transparency as key design concerns. This underscores the need to address unethical behavior in CPES design. A recent literature review of 56 studies [19] highlights *Privacy*, *Security*, and *Transparency* as the primary ethical requirements to be considered in the design of CPS. While [19] aims to identify and classify relevant ethical requirements, our work focuses on operationalizing these requirements through design techniques and evaluating them using a modeling framework. For this reason, we focus on privacy, security, and transparency, as they can be directly operationalized through design-level mechanisms such as data-access control, secure communication, traceability, and explanation support. We acknowledge that fairness and trust are also relevant

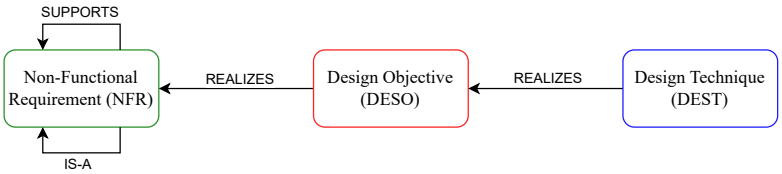

Fig. 1: Entity-Relationship Diagram of the Knowledge Graphs in DeTe4CPES.

ethical requirements for CPSs [19]. However, fairness requires additional criteria for comparing outcomes across stakeholders, tasks, or resources, while trust is typically shaped by multiple system properties, including privacy, security, and transparency. Therefore, we treat fairness and trust as important extensions to be addressed in future versions of the framework.

Summarizing, to the best of our knowledge, state-of-the-art methodologies lack conceptual approaches that properly structure and organize design solutions to address ethics-based concerns. To bridge this gap, we introduce DeTe4CPES whose novelty lies in outlining design objectives and techniques to address privacy, security, and transparency requirements. Business Processes are adopted as a target modeling formalism to illustrate ethical concerns in CPS design.

3 Preliminaries and Background

In this section, we provide the background for our approach. We begin with a concise overview of the European Parliament Scientific Foresight Study [10], followed by definitions of key terms used throughout this work.

European Parliament Scientific Foresight Study. The European Parliament [10] investigated the ethical implications of CPS, thereby offering insights into the potential impacts of their future evolution. Our paper builds on these considerations and focuses on understanding how moral principles have been incorporated into CPS design. A key issue is the trade-off between human oversight and machine autonomy, since CPS decide when to act, but humans may intervene, risking interference.

Knowledge Base. We adopt the conceptual model defined in [33] as the foundation of our work and enhance it with CPES-specific requirements by incorporating recent state-of-the-art design principles for their implementation. The extraction of concepts follows the definitions provided in [33], specifically: (i) *non-functional requirements* describe the target of a software system, e.g., privacy; (ii) *design objectives* represent the goals to realize the target, e.g., data protection; (iii) *design techniques* are concrete architectural or implementation-level solutions used to address the requirements, e.g., encryption. Rezayat et al. [32] also use the same model [33] and propose a framework for Hippocratic APIs that integrates GDPR principles into knowledge graphs.

Knowledge Graphs. We represent our approach as a set of knowledge graphs. Figure 1 shows the Entity-Relationship Diagram where each node represents an entity, such as non-functional requirement *(NFR)*, design objective *(DESO)*, or design technique *(DEST)*. Directed edges depict relationships and are labeled according to their type: `IS-A` denotes hierarchical specialization among *NFRs*; `REALIZES` indicates that a *DESO* addresses a *NFR*, or that a *DEST* implements a *DESO*; `SUPPORTS` links complementary NFRs that reinforce one another without implying subclassing, thereby enhancing the parent concept's effectiveness.

4 Shortcomings and Research Methodology

4.1 Ethics-Based Weaknesses in CPS Design

The Scientific Foresight project "Ethical Aspects of Cyber-Physical Systems" [10] was requested by the Science and Technology Options Assessment Panel (STOA)[1], and it offers valuable guidance. In the list of possible issues and challenges, it is clearly stated that in the context of CPS *new vulnerabilities may be exploited by hackers either to corrupt the operation of systems, or to extract commercial or other sensitive data* [10]. To avoid these vulnerabilities, it is necessary to incorporate the human-centered approach directly into the design phase. The following weaknesses of CPS design practice are addressed by our approach:

(W_1) Traditional CPS lack protection mechanisms to verify that **Privacy** is guaranteed, thus leading to significant vulnerabilities, e.g., there was the case of cleaning robots sharing images of people in bathrooms[2] . Moreover, the legal regulations may vary across Europe, e.g., vehicle tracking is prohibited in Germany, while permitted in France [10]. Hence, the monitoring of certain information (i.e., CPS data acquisition) must follow a legal path.

(W_2) Traditional CPS operate in dynamically changing environments, and data streams become ubiquitous and increasingly exposed to **Security** attacks that call for updating security measures constantly over time. A recent study estimated that cybercrime will have a tremendous monetary cost in the near future, approximately estimated to be nearly $14 trillion by 2028, and even data breaches will have an average cost of $4.45 million[3].

(W_3) CPS often lack **Transparency** in data management, making it difficult for humans to know how their data is collected, used, and shared. Two aspects need to be considered here: (i) transparency in the operation (e.g., a tele-presence robot requires control functions and safeguards); (ii) transparency in the process, (e.g., the terms of use of algorithms that rely on data for which the humans are not only aware of, but also approve their usage) [10]. Overall, an insufficient explanation of CPS behavior contributes to a lack of trust by humans.

[1] Panel for the Future of Science and Technology (STOA).
[2] A Roomba recorded a woman on the toilet. How did screenshots end up on Facebook?
[3] Industrial robots are at increased risk of cybersecurity threats.

4.2 Research Methodology

Here, we present our approach to capturing existing design techniques that address Privacy, Security, and Transparency in CPS. To guide our study, we formulate the following research question:

RQ: *What design techniques have been proposed and/or applied to address Privacy, Security, and Transparency requirements in the design of CPS?*

To answer this research question, our methodology consists of investigating the search string reported hereafter:

DESIGN *OR* DEVELOP* *OR* MANAGE* *OR* ARCHITECT* **AND** "CYBER-PHYSICAL SYSTEM" *OR* "CYBER PHYSICAL" **AND** "NON-FUNCTIONAL REQUIREMENT" **AND** PRIVACY *OR* SECURITY *OR* TRANSPARENCY

To ensure the selection of relevant studies, we define the following criteria:

- *Inclusion criteria:* the paper is written in English; it has been peer-reviewed; and it specifies design techniques on ethics only in the context of CPS.
- *Exclusion criteria:* the paper was published before 2020; it is a duplicate or a shorter version of an already included work; or it is off-topic, i.e., the search terms are used for other purposes.

The term "non-functional requirement" was included to restrict the search to studies that explicitly frame privacy, security, and transparency as requirements-level concerns, consistently with the scope of this work. We acknowledge that this may exclude relevant CPS design studies that do not use this terminology, but removing the term made the search too broad. Besides, we decided to restrict the search to papers published from 2020 onward to focus on the recent evolution of ethical concerns in CPS, thereby collecting contemporary design techniques towards operationalizing such requirements at design time.

The search process uses the ACM Digital Library, Scopus, and IEEE Xplore, as they are considered reliable sources by other literature reviews [17]. We focus on these databases because they index the main body of relevant peer-reviewed software engineering literature, including leading conference and journal venues.

Our replication package [18] includes all selected papers and maps each to the considered non-functional requirements, i.e., Privacy, Security, and Transparency. Overall, we consider 16 papers: 7 outline design strategies for addressing privacy weaknesses, 15 for security, and 3 for transparency. After selecting the papers, we performed a structured concept-extraction process. First, one researcher manually inspected each paper and extracted design-oriented concepts related to privacy, security, and transparency. The extracted concepts were then coded according to the ethical requirement they addressed and grouped by semantic similarity to identify recurring design strategies. To improve consistency,

the coding and grouping were reviewed by the other two authors, and disagreements were discussed until consensus was reached. The final agreed set of concepts was organized into the DESO and DEST knowledge graphs, following the structure described in Section 3. Note that some research works address multiple requirements, proposing design solutions that combine multiple mechanisms (e.g., architectural and cryptographic) that may jointly contribute to privacy, security, and transparency in the design of CPS. To provide an example of the selection rationale, [27] was considered to extract design strategies primarily related to *security*. The proposed approach focuses on the automated integration and verification of security requirements at design time through explicit threat modeling and mitigation. By reducing the risk of unauthorized access and security breaches in IoT-based services, this approach systematically satisfies security requirements. As another example, let us argue, as in [6], that the design strategies are suitable for all non-functional requirements considered. This is a survey on the adoption of fog computing platforms for smart city applications, and it reviews several design techniques, e.g., privacy can be ensured with encryption, security can be strengthened by decreasing the amount of data sent to a cloud infrastructure, and transparency can be improved with virtual objects that establish direct connections among sensors and actuators.

5 Design Approach for Cyber-Physical Ethical Systems

In this section, we present DeTe4CPES focusing on the non-functional requirements of Privacy, Security, and Transparency. For each requirement, we organize insights from our literature analysis into knowledge graphs that map DESign Objectives (DESO) to DESign Techniques (DEST). In the following, we describe DESO and DEST used in the use case (see Section 6), the complete description is provided in our replication package [18].

5.1 Design Objectives and Design Techniques for Privacy

Privacy is a foundational ethical principle that safeguards individual rights and ensures the respectful handling of sensitive information [10]. In current CPS deployments, continuous sensor data collection and processing intensify privacy challenges, making user control over personal data essential. Overall, for the Privacy requirement, we identified 7 design objectives and 12 corresponding design techniques, which are structured in the knowledge graph shown in Figure 2.

Security supports Privacy by preventing unauthorized access and restricting data handling to legitimate parties, thereby allowing individuals to maintain control of their personal information [25]. *Confidentiality* promotes Privacy by guaranteeing that sensitive data is accessible only to authorized entities.

DESO *Data Protection.* It focuses on preserving both data and model confidentiality by countering system-oriented attacks on sensitive information. This design objective encompasses the measures that prevent unauthorized access, use, disclosure, disruption, modification, or destruction of sensitive data.

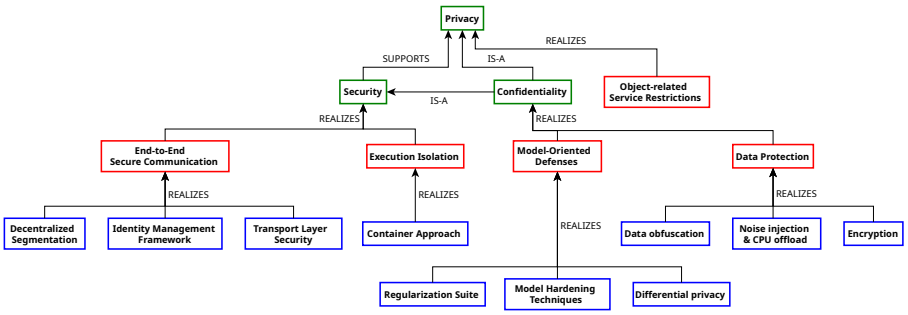

Fig. 2: Knowledge Graph mapping *Privacy* to its design objectives and their corresponding design techniques.

DEST *Data Obfuscation.* It transforms data to make the original content unreadable or less informative to attackers while remaining functional for authorized operations [8].
DEST *Encryption.* It converts readable data into an unreadable format (i.e., ciphertext) using cryptographic keys [37]. A key can only revert the ciphertext to its original form, ensuring that sensitive information remains confidential.

5.2 Design Objectives and Techniques for Security

Cyber-Physical Systems show the peculiar characteristic of integrating different components, such as sensors, actuators, and control algorithms. This contributes to vulnerability, since there might be attacks that can manipulate or disrupt critical operations in the absence of robust security measures. To address these risks, we identify and organize 17 design objectives and 31 design techniques related to security requirements. Since the resulting knowledge graph is extensive, we divide it into three parts for readability. Figure 3 presents the portion of the knowledge graph that includes the design objectives and techniques used in the use case discussed in Section 6. The remaining portions of the knowledge graph (along with DESO and DEST) are reported in [18].

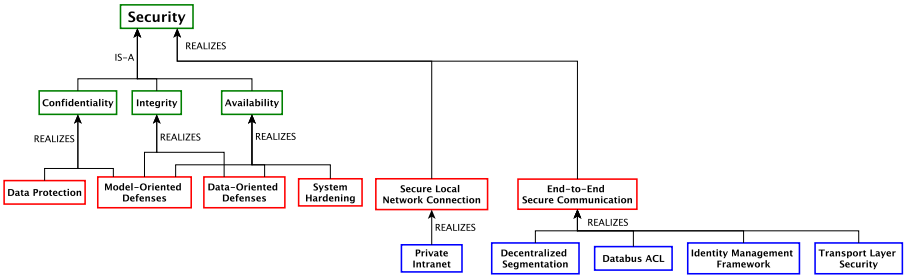

Fig. 3: An excerpt of the Knowledge Graph mapping *Security* to selected design objectives and techniques used in the use case presented in Section 6.

Confidentiality, *Integrity*, and *Availability* are foundational security NFRs [3]. Confidentiality prevents unauthorized disclosure of sensitive data; integrity preserves the completeness and trustworthiness of the data throughout its lifecycle; and availability ensures that services remain accessible as intended.

DESO *End-to-End Secure Communication.* It ensures data protection across all layers and intermediaries, regardless of the number of components involved.
DEST *Transport Layer Security.* It is a protocol that guarantees confidentiality, integrity, and authenticity of all communications between the components and fog nodes [6]. By establishing an encrypted and authenticated channel, it prevents the leakage and tampering of data during exchange.

DESO *Secure Local Network Connection.* It ensures that all communications between sensors, actuators, and local servers occur within a dedicated, isolated network, preventing external access and safeguarding data security [24]. Besides, positioning applications by differentiating geographic locations can foster local data processing and decrease the amount of information sent to cloud infrastructures, thus improving the security of restricted data [6].
DEST *Private Intranet.* It connects all device nodes to a local server through switches and routers in an intranet, creating an isolated backbone that protects internal traffic from unauthorized access [16].

5.3 Design Objectives and Techniques for Transparency

Transparency in CPS ensures that humans understand how system decisions are made and how these decisions affect system behavior. This is helpful for diagnosing faults, validating safety properties, and building users' trust. To support these needs, we extract and structure 6 design objectives and 5 design techniques, which are reported in Figure 4. The complete description of design objectives and techniques for transparency requirements is reported in [18].

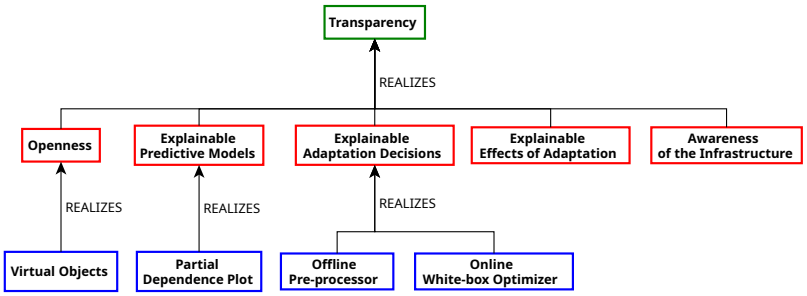

Fig. 4: Knowledge Graph mapping *Transparency* to its design objectives and their associated design techniques.

DESO *Openness.* It provides controlled access to system internals through well-defined interfaces, allowing agents and external services to observe and interact with core components. This improves transparency by making the system's functionality and state visible, understandable, and easier to monitor [35].

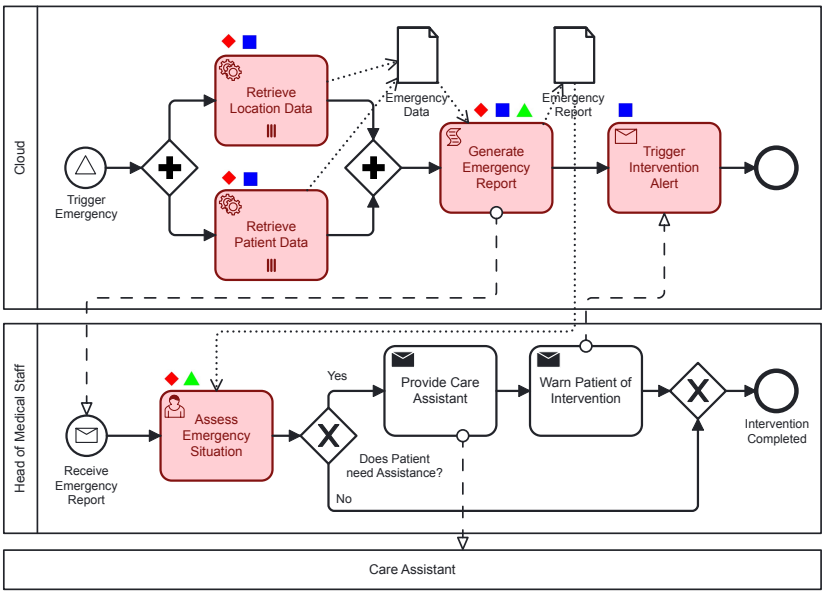

Fig. 5: BPMN Diagram of the Patient Monitoring Process. BPMN Activities highlighted in red indicate activities exposed to ethical weaknesses, graphically denoted by symbols: ◆ *Privacy*, ■ *Security*, ▲ *Transparency*.

DEST *Virtual Objects.* They are deployed above the physical sensors and actuators, and such objects expose standardized Application Programming Interfaces (APIs) [4,11]. These APIs provide transparent access to the underlying sensors and actuators, revealing the features and capabilities of physical objects. This empowers agents to establish direct connections to devices via APIs, eliminating the need for proprietary drivers or communication protocols [6].

6 Use Case: The Patient Monitoring System

To illustrate our approach, we map the shortcomings identified in Section 4.1 to our BPMN-based use case. We then discuss how these issues can be addressed by the appropriate DESO and their corresponding DEST outlined in Section 5.

It is worth noting that the conceptualization of ethics-based requirements, DESO, and DEST relies on extracting domain knowledge in the field of CPS and is intentionally not tied to any modeling formalism.

6.1 Business Process Model

We present the patient monitoring use case as a business process model to provide a structured representation of the scenario under analysis. A business process is a dynamic sequence of activities, events, and decisions that collectively

Table 1: Mapping of Ethical Weaknesses with BPMN Activities.

BPMN Activity	W_1 (◆ Privacy)	W_2 (■ Security)	W_3 (▲ Transparency)
Retrieve Location Data	✓ Patient profiling without anonymization	✓ Data are sent over unencrypted channels	—
Retrieve Patient Data	✓ Patient profiling without anonymization	✓ Data are sent over unencrypted channels	—
Generate Emergency Report	✓ Use of personal and medical data	✓ The report could be manipulated	✓ Patient is unaware of which data is used
Trigger Intervention Alert	—	✓ An attacker could trigger the intervention alert	—
Assess Emergency Situation	✓ Access to personal and medical data	—	✓ Patient is unaware of emergency strategies

produce or deliver a service [7]. Figure 5 depicts the workflow of the process using the well-known standard BPMN language [5,26].

Context. The use case concerns an assisted-living facility that comprises patients and medical staff (including a head who verifies emergencies and care assistants). Patients are equipped with a *wearable device* that monitors their locations and vital parameters. The process includes: (i) a **cloud** component using location and medical data of patients to generate the emergency report; (ii) the **head of medical staff**, who serves as the primary healthcare coordinator; (iii) the **care assistants** who are requested to intervene in case of emergency.

Process Workflow. The patient monitoring process begins when a *wearable device* detects anomalous vital parameters and triggers an alarm. The emergency data is processed by the underlying cloud infrastructure to generate a report, which is sent to the head of the medical staff. In cloud-based processing, the system combines live sensor data (e.g., location, heart rate) with patient-specific information (e.g., age, gender, prior health conditions). The head of the medical staff reviews the generated report and, when necessary, a care assistant is provided, and the patient is warned of the intervention.

6.2 Ethical Shortcomings

Table 1 reports the mapping of the ethical weaknesses with BPMN activities.

(W_1) **Privacy.** During the monitoring, the system collects highly sensitive, non-anonymized location data (*Retrieve Location Data*) and medical records, including patients' identities and clinical details such as age and treatments. Location data enables tracking and profiling, while clinical records contain personal information. When generating the Emergency Report (*Generate Emergency Report*) and assessing it (*Assess Emergency Situation*), identifiers are combined with medical data. This creates the trade-off between protecting patient privacy and ensuring an effective monitoring strategy that accounts for individual needs.

(W_2) **Security.** If communication between system components and the cloud is unencrypted or unauthenticated, attackers can intercept sensitive data (*Retrieve Location Data* and *Retrieve Patient Data*), causing data leaks. They may

Table 2: Mapping of Identified Weaknesses and their corresponding Design Objectives (DESO) and Design Techniques (DEST) in the Use Case.

Weakness	**Design Objective (DESO)**	**Design Technique (DEST)**
W_1 (◆ Privacy)	+ Data Protection	+ Encryption + Data Obfuscation
W_2 (■ Security)	+ End-to-End Secure Communication + Secure Local Network Connection	+ Transport Layer Security + Private Intranet
W_3 (▲ Transparency)	+ Openness	+ Virtual Objects

also inject malicious actuators commands, e.g., *Trigger Intervention Alert*. During report generation (*Generate Emergency Report*), attackers could access and alter critical data, such as misreporting patient needs, leading to unsafe care decisions. This creates a trade-off between preventing unauthorized data manipulation and ensuring timely emergency generation for prompt action.

(W_3) **Transparency.** After generating the emergency report (*Generate Emergency Report*), the head of medical staff assesses the situation (*Assess Emergency Situation*) and dispatches care assistants without revealing the prioritization rules derived from collected data. These rules explain the reason why some patients receive assistance before others. Providing access to such criteria could improve transparency and trust by enabling reconstruction of the decision-making process. This creates a trade-off between ensuring transparency in intervention strategy and protecting patients' private data.

6.3 Addressing Ethical Shortcomings

Table 2 maps the identified shortcomings to DESO and DEST.

(W_1) **Privacy.** We identify two primary privacy concerns (see Section 6.2): (i) the risk of patient profiling due to the use of non-anonymized sensitive information (i.e., location and medical data), and (ii) the uncontrolled use or access to such data during the generation of the emergency report. DeTe4CPES addresses these shortcomings by focusing on achieving *Data Protection (DESO)* to balance the use of critical information for emergency response with the need to safeguard patient privacy. While patient data must be used to support care assistant interventions, it should be transformed in a way that prevents patient identification. Techniques such as *Encryption (DEST)* or *Data Obfuscation (DEST)* can be applied to ensure that sensitive details are protected, while still allowing the necessary operational information to be accessible and understandable.

(W_2) **Security.** We identify three main concerns (see Section 6.2): (i) preventing the leakage of sensitive data, (ii) protecting against the malicious use of such data, and (iii) securing access to actuators (e.g., the generation of the emergency report) to avoid unauthorized control. DeTe4CPES offers multiple strategies (see Section 5.2) to address these security concerns. In the context of our use case, we propose adopting *End-to-End Secure Communication (DESO)* to prevent data leakage and *Secure Local Network Connection (DESO)* to isolate internal traffic from external threats. These mechanisms are particularly

relevant given the risks associated with the malicious manipulation of sensitive data and actuators (i.e., unauthorized generation of the emergency report). The proposed goals are achieved through specific Design Techniques: *Transport Layer Security (DEST)* to authenticate and encrypt all communication channels, and *Private Intranet (DEST)* to create isolated network segments, thus preventing external attacks and hindering unauthorized access. Hence, in this case, we prioritize *End-to-End Secure Communication (DESO)* and *Secure Local Network Connection (DESO)* because they directly mitigate sensor-stream tampering and command-injection threats at the communication and network boundaries.

(W_3) **Transparency.** We identify a key transparency issue (see Section 6.2): the lack of clarity in the prioritization rules used to determine which patients receive priority in assistance. DeTe4CPES proposes adopting the *Openness* (*DESO*) by applying the *Virtual Object* (*DEST*) technique, which exposes abstract representations of system components through standardized APIs. This enables stakeholders to review the data and device states behind the emergency report, aiding decision reconstruction and enhancing system accountability. *Openness*, realized through the *Virtual Object (DEST)*, was selected as the least intrusive DESO that still provides a meaningful level of auditability for this use case. By exposing standardized metadata and relevant runtime state through stable APIs, stakeholders can inspect the evidence underlying the emergency report and prioritization outcomes without direct access to sensitive raw identifiers or low-level infrastructure details. In the knowledge graph for Transparency (see Figure 4), this choice is made explicit via the `REALIZES` links from the Transparency NFR to *Openness* and then to *Virtual Object*.

6.4 Prototype Implementation of DeTe4CPES

To foster a quantitative evaluation of the design techniques, we provide a prototype implementation [18], i.e., a Python-based simulator that generates discrete-event traces of the activities and decisions in the BPMN model. Our prototype tool includes an interface (please refer to [18]) through which software designers can select the most suitable design techniques and assign numerical values to indicate their impact. Each DEST can be applied at different intensity levels, e.g., Low (L), Medium (M), and High (H), which represent increasingly stronger implementations of the corresponding technique, e.g., the key size of the adopted encryption algorithm. For instance, our use case includes Privacy endorsed with the following DESTs: *Encryption* and *Data Obfuscation* (see Section 6.3). These DESTs can be implemented by invoking libraries[4] that provide concrete transformations of the data exchanged during the simulated process. *Encryption* encodes sensor and actuator payloads before transmission, while *Data Obfuscation* pseudonymizes the patient's personal details and perturbs the transmitted data.

To use the prototype implementation, software designers must define system parameters, i.e., the number of monitored patients (e.g., 50) within an observation window (e.g., a 30-day simulation horizon). Moreover, software designers can

[4] Sources for Encryption, and Data Obfuscation.

specify the formulas for calculating *Privacy*, *Security*, and *Transparency* scores, thereby determining how best to combine the selected DESTs. This way, the quantitative evaluation can be tailored to designers' specific needs. The prototype is intended as an illustrative proof of concept showing how DeTe4CPES can be made actionable and configured by designers. Its empirical validation, e.g., studies with software designers to assess its usability, is part of future work.

7 Threats to Validity

Internal validity. The mapping between non-functional requirements, design objectives, and techniques may be affected by researchers' interpretations when extracting concepts from the literature. Besides, the presented knowledge graphs cover the design techniques found in the literature but may neglect other aspects, such as the socio-technical, legal, and human-centered dimensions of these requirements. To mitigate these threats, the mappings rely on explicit NFRs (i.e., privacy, security, transparency), and we make the knowledge graphs and extraction artifacts publicly available [18], enabling future extension.

Construct validity. The quantitative evaluation of design techniques, along with the specification of *Privacy*, *Security*, and *Transparency* scores, is delegated to software designers. To address this threat, we plan to integrate a system component that monitors real deployment data and quantifies the actual impact of adopting different DESTs, both in isolation and in combination.

External validity. A potential concern is the scalability of knowledge graphs. While they are promising for structuring non-functional requirements, design objectives, and techniques, their use in highly complex, multi-domain CPS may increase modeling effort and reduce maintainability. To mitigate this threat, DeTe4CPES is defined independently of a specific modeling formalism, and the identified concepts are organized in reusable knowledge graphs. A further threat concerns the illustrative nature of the use case, which shows how DeTe4CPES can be applied but gives no evidence of generalization that is left as future work.

8 Conclusion

This paper introduces DeTe4CPES, a novel conceptual approach for designing Cyber-Physical Ethical Systems (CPES). It focuses on three core ethics-based weaknesses in CPS: *Privacy*, *Security*, and *Transparency*. We study the ethical challenges identified in the literature and structure the proposed design objectives and techniques as knowledge graphs. In addition to addressing the threats to validity, our future work includes the following main activities. We plan to model design techniques and study their application in different CPS scenarios. We also aim to operationalize these techniques by monitoring metrics that quantify their effects, thereby automating the quantitative evaluation of CPES.

Acknowledgments. We thank the anonymous reviewers for their valuable feedback. This work has been partially funded by the MUR-PRIN project 20228FT78M DREAM and by the MUR Department of Excellence 2023–2027 for GSSI.

References

1. Borchert, A., Ferreyra, N.E.D., Heisel, M.: A conceptual method for eliciting trust-related software features for computer-mediated introduction. In: Int. Conf. on Evaluation of Novel Approaches to Software Engineering. pp. 269–280 (2020)
2. Carreño-Bolufer, J., et al.: A conceptual model-based application for the treatment and management of data in pediatric oncology: The neuroblastoma use case. In: Int. Conf. on Evaluation of Novel Approaches to Software Engineering. pp. 15–26 (2025)
3. Chen, H., Babar, A.: Security for machine learning-based software systems: A survey of threats, practices, and challenges. ACM Comput. Surv. **56**(6), 1–38 (2024)
4. Cicirelli, F., Guerrieri, A., Spezzano, G., Vinci, A.: An edge-based platform for dynamic smart city applications. Future Gener. Comput. Syst. **76**, 106–118 (2017)
5. Compagnucci, I., Corradini, F., Fornari, F., Re, B.: A Study on the Usage of the BPMN Notation for Designing Process Collaboration, Choreography, and Conversation Models. Business & Information Systems Engineering **66**, 43–66 (2024)
6. Da Silva, T.P., Batista, T., Lopes, F., Neto, A.R., Delicato, F.C., Pires, P.F., Da Rocha, A.R.: Fog computing platforms for smart city applications: A survey. ACM Transactions on Internet Technology **22**(4), 1–32 (2022)
7. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management (2018)
8. Enireddy, V., Somasundaram, K., Prabhu, M.R., Babu, D.V., et al.: Data obfuscation technique in cloud security. In: Int. Conf. on Smart Electronics and Communication (ICOSEC). pp. 358–362 (2021)
9. España, S., van der Maaten, C., Gulden, J., Pastor, Ó.: A survey of ethical reasoning methods, their metamodels, and a discussion on their application to conceptual modelling. In: Int. Conf. on Conceptual Modeling (ER). pp. 23–44 (2023)
10. European Parliamentary Research Service: Scientific Foresight - Ethical Aspects of Cyber-Physical Systems. Link - Last Access: 15 May 2026 (2016)
11. Giordano, A., Spezzano, G., Vinci, A.: Smart agents and fog computing for smart city applications. In: Int. Conf. on Smart Cities. pp. 137–146 (2016)
12. Guizzardi, R., Amaral, G., Guizzardi, G., Mylopoulos, J.: An Ontology-based Approach to Engineering Ethicality Requirements. Software and Systems Modeling **22**(6), 1897–1923 (2023)
13. Hästbacka, D., Halme, J., et al.: Dynamic edge and cloud service integration for industrial iot and production monitoring applications of industrial cyber-physical systems. IEEE Trans. Ind. Informatics **18**(1), 498–508 (2022)
14. He, X., Dong, Z., Yin, H., Fu, Y.: A Framework for Developing Cyber-Physical Systems. Int. J. Softw. Eng. Knowl. Eng. **27**(9-10), 1361–1386 (2017)
15. Hehenberger, P., Vogel-Heuser, B., Bradley, D., Eynard, B., Tomiyama, T., Achiche, S.: Design, modelling, simulation and integration of cyber physical systems: Methods and applications. Computers in Industry **82**, 273–289 (2016)
16. Helal, R., Seghiri, A., Belala, F., Hameurlain, N.: Towards a Formal Modeling Approach for Cyber-Physical Systems Requirements. In: Int. Conf. on Software and Computer Applications (ICSCA). pp. 298–309 (2024)
17. Kitchenham, B.A., et al.: Systematic literature reviews in software engineering - A systematic literature review. Inf. Softw. Technol. **51**(1), 7–15 (2009)
18. Kurkchi, Y., Compagnucci, I., Trubiani, C.: Replication Package – "Understanding the Concepts behind Ethics in Designing Cyber-Physical Systems" (2026), Link

19. Kurkchi, Y., Trubiani, C.: Ethics-Based Requirements for Engineering Cyber-Physical Systems: A Literature Study. In: Euromicro Conference Series on Software Engineering and Advanced Applications (SEAA). pp. 254–272 (2025)
20. Lee, E.A.: Cyber Physical Systems: Design Challenges. In: Symposium on Object and Component-Oriented Real-Time Distributed Computing. pp. 363–369 (2008)
21. Lee, J., Bagheri, B., Kao, H.A.: A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manufacturing letters **3**, 18–23 (2015)
22. Leitão, P., Karnouskos, S.: The Emergence of Ethics Engineering in Industrial Cyber-Physical Systems. In: Int. Conf. on Industrial Cyber-Physical Systems (ICPS). pp. 1–6 (2022)
23. Lukyanenko, R., Samuel, B.M., Parsons, J., Storey, V.C., Pastor, O., Jabbari, A.: Universal Conceptual Modeling: Principles, Benefits, and an Agenda for Conceptual Modeling Research. Softw. Syst. Model. **23**(5), 1077–1100 (2024)
24. Maruf, M.A., Azim, A.: Automated features and requirements identification for improving cps software reuse using topic modeling. In: Int. Conf. on Cyber-Physical Systems (ICCPS). pp. 262–263 (2023)
25. Milossi, M., Alexandropoulou-Egyptiadou, E., Psannis, K.E.: AI ethics: algorithmic determinism or self-determination? The GPDR approach. IEEE Access **9**, 58455–58466 (2021)
26. OMG: Business Process Model and Notation (BPMN), Version 2.0.2. Normative Document - Last Access: 15 May 2026 (2013)
27. Ooi, S.E., Beuran, R., Tan, Y., Kuroda, T., Kuwahara, T., Fujita, N.: Secureweaver: Intent-driven secure system designer. In: Int. Workshop on Secure and Trustworthy Cyber-Physical Systems (SaT-CPS). pp. 107–116 (2022)
28. Paja, E., Dalpiaz, F., Giorgini, P.: Modelling and reasoning about security requirements in socio-technical systems. Data & Knowledge Engineering **98**, 123–143 (2015)
29. Prakash, D., Prakash, N.: A conceptual model for data warehousing. In: Int. Conf. on Evaluation of Novel Approaches to Software Engineering. pp. 87–98 (2024)
30. Prakash, N., Prakash, D.: Concepts for conceptual modelling of an iot application. In: Int. Conf. on Evaluation of Novel Approaches to Software Engineering. pp. 494–501 (2022)
31. Rehman, S.u., Allgaier, C., Gruhn, V.: Security Requirements Engineering: A Framework for Cyber-Physical Systems. In: Int. Conf. on Frontiers of Information Technology (FIT). pp. 315–320 (2018)
32. Rezayat, S., Burmester, G., Ma, H., Hartmann, S.: Conceptual Framework for Designing Hippocratic APIs. In: Int. Conf. on Conceptual Modeling (ER). pp. 355–376. Springer (2024)
33. Sadi, M.H., Yu, E.: WEBAPIK: a Body of Structured Knowledge on Designing Web APIs. Requir. Eng. **28**(3), 441–479 (2023)
34. Tantawy, A., Abdelwahed, S., Erradi, A.: Cyber LOPA: An Integrated Approach for the Design of Dependable and Secure Cyber-Physical Systems. IEEE Transactions on Reliability **71**(2), 1075–1091 (2022)
35. Torresen, J., Saplacan, D., Baselizadeh, A., Mahler, T.: Machine excellence trade-offs to ethical and legal perspectives. In: Int. Conf. on Artificial Intelligence (CAI). pp. 237–240 (2023)
36. Zhang, Q., Mohammed, A.Z., Wan, Z., Cho, J.H., Moore, T.J.: Diversity-by-Design for Dependable and Secure Cyber-Physical Systems: A Survey. IEEE Transactions on Network and Service Management **19**(1), 706–728 (2022)
37. Zhu, Y., Cheng, Y., Zhou, H., Lu, Y.: Hermes attack: Steal DNN models with lossless inference accuracy. In: USENIX Security Symposium. pp. 1973–1988 (2021)