Implications of Trust in Cyber-Physical Systems Design: the ASSA Case Study

Pierre Rambert[1] and Irina Rychkova[1]

University Paris 1, Panthèon-Sorbonne, Paris, France
{pierre.rambert,irina.rychkova}@univ-paris1.fr

Abstract. Cyber-Physical Systems (CPS) refer to integrated computational and physical processes, where computational elements monitor and control physical processes, usually with a feedback loop.Analysis of trust formation within a complex networks of linked social, physical and computational entities has a potential to reveal otherwise implicit trustworthiness requirements and to improve the CPS design and acceptance. We address this challenge with a method for trust analysis that supports the alignment between trust concerns expressed by the prospective CPS users and stakeholders, trustworthiness requirements formulated by the CPS designers and technical elements of the CPS to be developed. We apply this method to a real CPS design project. Our case study shows that the explicit analysis of trust not only improves traceability between user trust concerns and technical features of the CPS, but also allows for identification of new relevant trustworthiness requirements, affecting the system design and acceptance.

Keywords: trust, trustworthiness requirements, CPS

1 Introduction

Cyber-physical systems (CPS) refer to integrated computational capabilities, networking, and physical processes. These systems use embedded computers and networks to monitor and control physical processes, often with feedback loops, where physical actions influence computational decisions and vice versa [12]. Social entities (organizations and individuals) are inherently involved in the CPS lifecycle from development and standardization to implementation and daily use.

Trust is an essential component of the CPS adoption [6]. CPS are characterized by their high degree of complexity, interconnectivity, and the ability to interact with both the physical world and computational elements seamlessly. Compared to social or interpersonal trust [5,13], in a CPS context, there are multiple entities (social, physical, or computational) upon which CPS users need to place their trust. Examining trust formation in such a complex heterogeneous system provides valuable input for CPS designers and developers, improving systems' trustworthiness and positively affecting their acceptance and adoption by the users [6, 19, 24].

We address this challenge with a method for iterative analysis and elicitation of trustworthiness requirements elaborated from [16]. The method is grounded on the Six-Variable Model [25] originally defined to support design of control

systems. It provides a structured framework for presentation and analysis of the relationships between various CPS elements. The presented method supports traceability and alignment between *trust concerns* expressed by the CPS users and stakeholders, *trust assumptions* made by the CPS designers in order to address these concerns, *trustworthiness requirements*, and technical system *components* that will be developed to meet these requirements. This article presents the details of the method and reports on the case study of the ASSA project where this method was applied. ASSA (Assistance Sécurité Seniors Application) is a startup creating a personal emergency response system for the elderly. ASSA solution uses connected smart devices to monitor user's vital parameters and to trigger an alert in case of emergency.The team of two engineers delivered design documentation for ASSA following a conventional software design process. However, explicit trust analysis and trustworthiness requirements elicitation were not conducted. This led us to consider ASSA an appropriate case for the method application. The goal of our study is to demonstrate that explicit trust analysis offers valuable insights for the CPS design process, enhancing traceability and leading to the identification of new relevant requirements that contribute to the system's trustworthiness.

This article is organized as follows: In Section 2, we discuss the background on trust in CPS; In Section 3, we provide the details on the method of trustworthiness requirements elicitation; In Section 4, we present our research methodology and introduce the ASSA case study; In Section 5, we present our case study results; We provide the concluding remarks in Section 6.

2 Background

2.1 Trust in CPS Research

Trust is a social construct that emerges from interactions between individuals or groups and can be described by a situation where a subject (trustor) is willing to rely on a chosen actions of an object of trust (trustee) [20] [5] [13]. Advances in socio-technical systems introduce novel models of social and business interactions, where IT artifacts can take the role of a trustee [22]. Trust in technology reflects trustor's beliefs that a specific technology has the attributes necessary to perform as expected in a given situation where negative consequences are possible [14] [15].

CPS applications for assisted living provide individuals with support within their living environments [23] [4] [1]. Recent technological advances expanded the capabilities of CPS, enabling real-time health monitoring, fall detection, medication management, and personalized assistance. However, multiple issues related to privacy, data security, interoperability, standardization, and integration of CPS systems into larger ecosystems undermine the CPS users' and stakeholders' decision to trust and to be engaged with a CPS [6] [1] [19]. Trust, defined as the user's willingness to rely on a system in case of an emergency, is a prerequisite for CPS technology adoption and must be explicitly addressed in their design [24] [19] [17].

While interpersonal or social trust can be defined as a function of trustee's perceived ability, benevolence and integrity [13], physical and computational

entities exhibit technical properties and attributes that might predict the user's decision to trust (and by consequence to be involved with) the CPS [14].

Given the complex and interconnected nature of a CPS context, where multiple social, physical and computational entities involved, there may be no obvious central or identifiable trustee upon which to base trust decisions [6]. Moreover, non-technical CPS stakeholders have limited capacity to objectively assess technical properties of a CPS and to reason about its trustworthiness. In [6], the following eight trust constructs in CPS are defined: familiarity and understanding of the CPS by consumers; reliability, predictability and consistency; security; integrity; competence, expertise and functionality required to interact with the CPS; the benevolence and helpfulness of the CPS for consumers; personalizability; faith and belief consumers have in the service delivered. These concepts capture the complex nature of trust relationships in CPS. To accurately address CPS trustworthiness during design, it is essential to conduct a thorough analysis of user trust-related concerns and systematically translate these concerns into trustworthiness requirements [16].

2.2 Trustworthiness Requirements and Trust Assumptions

In systems engineering, trustworthiness of a system means "to be worthy of being trusted" to fulfill some specific requirements [18]. In this work, we elaborate on the method for explicit trustworthiness requirements elicitation proposed in [16]. Trustworthiness requirement (TwR) can be defined as *a statement made by a trustor about the expected trustworthiness of a trustee. This statement must clearly express an operational, functional, design, or other characteristic, which, according to the trustor's beliefs, positively impacts the trustworthiness of this trustee and the interaction between the two.*

Sutcliffe [24] introduces 'soft' requirements as a linguistic concept that encompasses various phenomena related to people, organizations, and society, including trust. Grounded on this work, TwR can be considered as a subclass of soft requirements and can be refined by functional and non-functional requirements. A RE method aiming to systematize the elicitation and analysis of requirements, including trustworthiness requirements, and grounded on the ontological analysis is proposed in [3]. Here, TwR are considered as a special class of quality requirements. The Reference Ontology of Trustworthiness Requirements (ROTwR) [2] proposes decomposition of TwR into reliability requirements, truthful information communication requirements and transparency requirements.

In the context of CPS, TwR define the desired outcomes these systems should achieve to address the trust concerns (TC) of users and stakeholders within the socio-physical environment. However, the CPS controlling software - the main focus of the development project addressed in this study - can only affect the machine domains, not the socio-physical ones. This concept is known as the world-on-machine paradox, as formulated in [26]. Consequently, TCs cannot be directly mapped into functional, non-functional, or quality requirements, nor can they be linked directly to CPS components within the machine domain. Instead, system engineers must interpret these concerns from the social domain to the machine domain [27] [7] [21]. This process involves delegating the user's trust decisions to system engineers, who make explicit or implicit choices to trust

certain technical characteristics or components. These choices are referred to as design assumptions in [9].

An ontology of assumptions proposed in [26] defines world assumptions (WA), machine assumptions (MA), world dependence assumptions (WDA), and machine dependence assumptions (MDA). A world assumption is an assumption about world (or social) phenomena, which constraints the machine environment. A machine assumption is an assumption about a machine's internal phenomena. A machine dependence assumption states that an external world phenomenon depends on some machine phenomena. In contrast, a world dependence assumption states that a machine phenomenon depends on some world phenomena. The authors use situation calculus to reason about assumptions and requirements.

Grounded on [8] [26], trust assumptions (TA) can be defined as the *assumptions made by the system engineers about the properties of a system-to-be and its various components (including human components), which will positively affect the perceived trustworthiness of the system.*

3 Method for Trustworthiness Requirements Elicitation

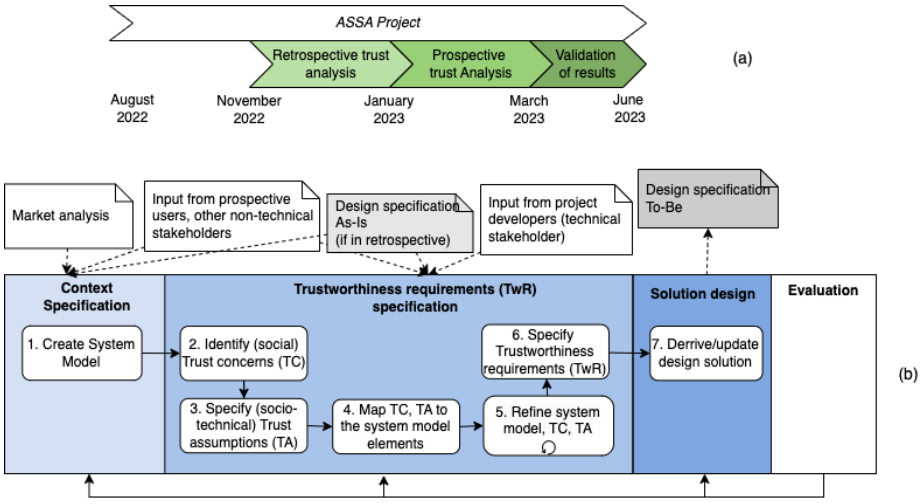

Fig. 1. (a) Overview of the ASSA Case Study: the timeline; (b) Method for Trust Analysis and TwR elicitation integrated into the human-centered design approach [10]

The method presented in this work is grounded on [16]. For trust analysis, we use the Six-Variable Model [25] originally defined to support design of control systems. This model provides a structured framework for analysis of the relationships between various system elements situated in the socio-physical environment (i.e., users, stakeholders) and in the machine domain (i.e., a control machine,

sensors, actuators, other connected systems). It defines six variables that are depicted as relations between different system elements: referenced variables that focus on the properties that should be observed in the system's environment; monitored variables represent the properties that need to be monitored during the system's operation; input variables represent the external stimuli that affect the system's behavior; output variables refer to the outcomes produced by the system; controlled variables refer to the properties that the system can actively control or actions the system can take to achieve its desired behavior; desired variables refer the properties in the system's environment that should be achieved during the system's operation. Figure 2 illustrates the six-variable model for the ASSA system (our case study).

By analyzing the relationships between monitored, controlled, input, output, referenced, and desired variables, developers can precisely articulate the system requirements – TwR in our case - and their implication of system functionalities and behavior. For trust analysis in CPS, we use the formalism proposed in [8] [16] for documenting TC, TA, TwR and traceability between TA, TwR and system elements.

Our method defines seven steps (see Fig. 1-b) as follows:
Step 1: A global system model based on the Six-Variable Model is created. This model focuses on the problem domain and the effects that must be achieved in this domain. It captures the context of use and forms the foundation for the subsequent steps.
Step 2: Trust concerns are collected from the users and project stakeholders. In case of retrospective analysis, trust concerns can be extracted from the existing design documentation;
Step 3: Trust assumptions are collected from the technical stakeholders. An assumption is *something taken as being true or factual and used as a starting point for a course of action or reasoning*[1]. Trust assumptions refer to the (social) trust concerns and justify the design choices made during the solution design. We use the four kinds of assumptions defined in [26] to address the 'world-on-machine paradox' and to align the problem domain and the machine domain.
Step 4: The trust concerns and trust assumptions are mapped to their relevant elements in the global system model. Each trust concern may refer to one or several elements in the system model.
Step 5: The global system model is refined: the high-level problem is refined into sub-problems. Trust concerns and trust assumptions are elaborated. If some trust concern is not addressed - a design assumption should be made and a new element, feature or property needs to be suggested for the solution. This step repeats until all trust concerns are covered and a desired level of detail is achieved;
Step 6: The trustworthiness requirements are derived from the trust assumptions and integrated into system design documentation.
Step 7: The new technical features and/or components are derived, contributing into trustworthiness of the design solution. The design specification is updated.

The method is consistent with the activities of the ISO 9241 human-centered design approach [10]: Specifying the context of use (Step 1); Specifying the user

[1] https://www.merriam-webster.com/thesaurus/assumption

requirements (Step 2-6); Producing the design solutions (Step 7). Based on the evaluation results, a new iteration of trust analysis and TwR elicitation can be triggered. The User-centered evaluation of the design is out of scope for this article and will be addressed in future.

4 The Case Study

To demonstrate our method for eliciting trustworthiness requirements and to evaluate its effectiveness, we conducted a case study on the ASSA software development project.

4.1 ASSA: the Application for Assistance and Security for Elderly

ASSA is an innovative personal emergency response application designed for iOS and Android devices. Potential users of ASSA are elderly people and people living in isolation, both geographically and socially. The application can potentially integrate with compatible smartwatches and other wearable devices to monitor vital parameters (e.g., heart rate, blood pressure, oxygen saturation levels) and detect accidents (e.g. falls). When an irregularity in vital parameters or an accident is detected, the application initiates a series of actions to ensure quick assistance. First, it triggers a timer, during which the user is prompted to confirm or cancel the emergency. If the emergency is not canceled, ASSA transmits the emergency alert to the designated healthcare providers (e.g., a hospital emergency service) and notifies a designated caregiver (e.g., a person of trust or a family member), providing them with the data related to the situation (e.g., a health report).

The team of two engineers conducted the market analysis, requirements specification and developed initial design documentation for ASSA between August and November 2022 (see Fig. 1-a). Regular monitoring and emergency handling are the main uses-cases of ASSA. Other functionalities have been elaborated later during the project. In particular, systematic generation and management of on-line health reports have been added to ASSA as a result of this study.

According to the latest design specification, the ASSA application ensures user monitoring and alert generation in case of emergency and provides relevant historical data for the medical professionals (for both regular and emergent medical interventions) for more efficient personalized treatment.

4.2 Research Methodology

We adapt the single-project case study research protocol defined by Kitchenham and Pickard in [11].

Planning and designing of the case study. The objectives of the case study is to evaluate the feasibility, efficiency and relevance of our method for the analysis and elicitation of trustworthiness requirements in the ASSA project. While trust was recognized by the ASSA developers as an important factor for adoption, no specific trust analysis has been conducted during the project. This made

the project a relevant case for the study. We define the following hypothesis for this case study:
H1: The method is complementary with a design process not focused on trust.
H2: The method application uncovers implicit TwR in the existing system specifications.
H3: The method application leads to identification of the new TwR (extending the system specifications).
H4: Documentation of trust assumptions enhances traceability and alignment in the designed system.
H5: Method application contributes into (re)definition of a valid and relevant design solution.

Conducting the case study. The study was conducted between November 2022 and June 2023 (Fig. 1-a). First, we conducted the trust analysis of the system *in retrospective*, applying the method on the data collected in the ASSA project before November 2022 and to the design documentation produced by the ASSA developers. In the second iteration, we conducted the trust analysis of the system *in prospective*, applying the method on the new data. We conducted seven semi-structured interviews with prospective ASSA users and healthcare professionals. Additionally, we organized a series of working sessions with the ASSA developers, concentrating on the themes of trust, trustworthiness, and acceptance of the prospective application. Each interview lasted between 20 and 45 minutes. After a brief presentation of the assisted living technologies, the interviewees have been introduced to ASSA and invited to discuss their own experience with the assisted living technology and their reasons to adopt (or not) one. Compared to interviews conducted by ASSA developers in the earlier days of the project, this data collection focuses on subjective trust issues and beliefs of the (non-technical) stakeholders. Here are some example questions from the interview guide (translated from French):
- How would the personal emergency response system improve your work or your personal security?
- For what reasons would you take a personal emergency response system?
- Do you have any preoccupation when a relative of yours uses a personal emergency response system?
- Does a "connected" personal emergency response system changes any preoccupation you have regarding a typical personal emergency response system? Both retrospective and prospective analysis demonstrated the method feasibility (H1). We addressed the method efficiency (H2, H3) by eliciting the TwR in both iterations. We were able to identify new features contributing to the system trustworthiness and to propose an update for the base line ASSA design solution (As-Is). To evaluate the method relevance (H4, H5), we presented the refined ASSA system model and the updated design specification to the ASSA developers for review and validation. The case study results are presented in the following session.

5 Trustworthiness Requirements Elicitation for ASSA

5.1 The Retrospective Trust Analysis

Step 1: The global system model of ASSA is created based on the market analysis and design documentation produced by the ASSA developers (Fig. 2). Here ASSA mobile application is represented as a control system in a Six-Value-Model [25]. The root requirement is defined by the main use case of ASSA: Provide monitoring and assistance. The problem domain includes the user (an elderly person whose health is monitored), her designated as caregivers and healthcare practitioners. The control machine or the software to-be represents the ASSA mobile application. The connection domain between the machine and the problem domains includes sensors (e.g., smartphone, smart watch) and the external systems with which the control machine interacts by sending alerts, reports, and emergency calls (e.g., health emergency services, messaging service, and online reporting system). Reference variables include "health report" and

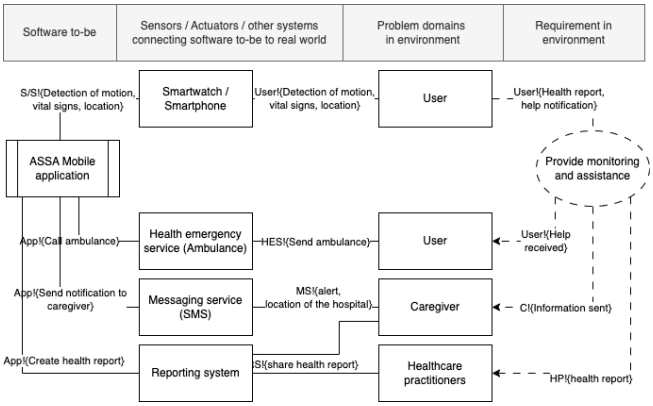

Fig. 2. Step 1: The ASSA global system model

"help notification"; monitored and input variables include "detection of motion", "vital signs" and "location"; output variables include "call ambulance" linked to the health emergency system (e.g., Ambulance), "send notification to caregiver" linked to the messaging service (e.g., SMS), and "create health report" linked to the reporting system. The controlled variables include "send ambulance" linked with the User; "alert", and "location of the hospital" linked to the Caregiver; "share health report" linked to the Healthcare practitioners and Caregiver. The desired variables are composed of "help received", linked to the User; "information sent", linked to the Caregiver; "health report", linked to the Healthcare practitioner.

Step 2: We conducted a semantic analysis of the ASSA design documentation (As-Is) provided by the ASSA developers, including the market reports and interviews with medical professionals, potential users and caregivers. Using various

definitions of trust from the literature, we identified seven trust concerns TC1-7 (Table 1).

Step 3: We identified the design decisions related to the TC specified in the previous step from the ASSA design documentation (As-Is). We identify the assumptions made by the ASSA developers about the properties of the ASSA CPS and its various components to justify these decisions. We conducted the interviews with the ASSA developers to validate the TA that have been discovered. We classify them according to the ontology proposed in [26]. Table 2 illustrates the list of TA with their related TC. During the retrospective trust analysis, we extracted 11 trust assumptions TA1.01-TA1.11. We use the assumption ontol-

Table 1. Trust concerns expressed by the prospective ASSA users and stakeholders

Iter.		Step 2: Trust concern
1	TC1	User is concerned about whether she will be constantly monitored
1	TC2	User is uncertain whether she will receive help upon feeling unwell
1	TC3	User is concerned with who will be able to access the monitoring data
1	TC4	User is concerned about whether she will be able to use the system in a proper way
1	TC5	User and caregiver are concerned about whether the device is charged and working
1	TC6	Caregiver is concerned about whether the assistance will be timely provided in response to the alert
1	TC7	Healthcare practitioners are concerned about relevance and exactitude of the health report
2	TC8	User is concerned about the system generating false alerts
2	TC9	User is concerned that the monitoring data and the health report will not be used by the healthcare professionals to personalize/improve the treatment
2	TC10	User is concerned that his data will be used for commercial purposes
2	TC11	User is concerned about the health report being shared/used without her consent

ogy from [26], identifying world assumptions (WA), machine assumptions (MA), world dependence assumptions (WDA), and machine dependence assumptions (MDA).

World assumptions are engineers assumptions about the social phenomena. For example, we assume that the healthcare provider sends an ambulance each time the ASSA emergency alert is received (see TA1.10) - this assumption means that our solution cannot be trusted if the healthcare provider ignores the alert or has not enough resources to respond to it. We also assume that the user should find the ASSA interface simple and intuitive (TA1.08). Otherwise the system will not be trusted by the user.

Machine assumptions are engineers assumptions about the internal properties of the application, smart devices or other (external) systems and services that

Table 2. Trust assumptions discovered from the retrospective analysis (TA1.xx) and made during the prospective analysis (TA2.xx) linking the trust concerns with system elements

	Trust Assumption:	Kind [26]	TC:	Linked to System elements:
TA1.01	Redundant sensors increase the monitoring reliability	MA	TC1	Smartwatch/Smartphone
TA1.02	User can check the sensor's status and activity	MDA	TC1	ASSA Mobile application; Smartwatch/Smartphone
TA1.03	Irregularities in health metrics are correctly identified from the data	WDA	TC2, TC8	ASSA Mobile application
TA1.03.01	User can trigger an alert manually	WDA	TC2, TC8	ASSA Mobile application
TA1.04	The alert transmission is guaranteed by the system	MDA	TC2	ASSA Mobile application; Network provider
TA1.04.01	WiFi and Mobile communication is reliable	MA	TC2	Network provider
TA1.05	Emergency calls are responded 24/7	WA	TC2	Health Emergency service
TA1.06	Data is encrypted	MA	TC3	ASSA Mobile application; Messaging service; Reporting system
TA1.07	Data can be accessed only by authorized persons with the verified identity	WDA	TC3, TC10-11	Messaging service, Reporting system, Caregiver, Healthcare practitioner
TA1.08	The interface is simple and intuitive	WA	TC4	ASSA Mobile application; Smartwatch/Smartphone
TA1.09	User is informed when the battery of the smartphone or the smartwatch is low	MDA	TC5	ASSA Mobile application
TA1.10	The healthcare provider handles emergency alerts received from the system	WA	TC6, TC9	Health emergency service, Healthcare authorities
TA1.11	Software is certified and recognised by healthcare authorities	WA	TC7, TC9	Healthcare authorities
TA2.01	Redundant sensors prevent from false alerts	MDA	TC8	Smartwatch/Smartphone
TA2.02	Explicit request for the user consent	MDA	TC10-11	ASSA mobile application, Reporting system
TA2.03	GDPR-complience	MDA	TC10-11	ASSA mobile application, Reporting system
TA2.04	Reporting system is highly available	MA	TC9	Reporting system, Healthcare practitioner
TA2.05	Healthcare practitioners recognise and use the health report	WA	TC9	Reporting system, Healthcare practitioner
TA2.06	Healthcare practitioners provide qualified help	WA	TC9	Healthcare practitioner
TA2.07	User training materials are provided	MDA	TC1, TC4-5	ASSA Mobile application; Smartwatch/Smartphone

will positively affect the perceived trustworthiness of the system. We assume that the redundant sensors in the system affect reliability of the patient monitoring (TA1.01).

A machine dependence assumption states that an external world phenomenon depends on some machine phenomena. For example, we assume that the user can trust the system if she can check and make sure that the sensor is working and constantly monitoring (see TA1.02).

A world dependence assumption states that a machine phenomenon depends on some world phenomena. For example, we assume that the system detecting an anomaly (and/or raising an alert) means that the user is not well (TA1.03, TA1.03.01).

Step 4: We map the TC (Step 2) and TA (Step 3) to the relevant parts of the ASSA system. We update the system model to show the traceability (Fig. 3). ç Concerns TC1, TC4, TC5 are related to ASSA mobile application and/or the smart devices - the User interface. TC2 questions the whole system and its purpose - it is related to the main requirement. TC3, TC7 are related to the ASSA interfaces for data exchange with the external systems. TC6 is related to the interfaces between the system and the control domain. They refer to service-level agreements and operational-level agreements that need to be defined.

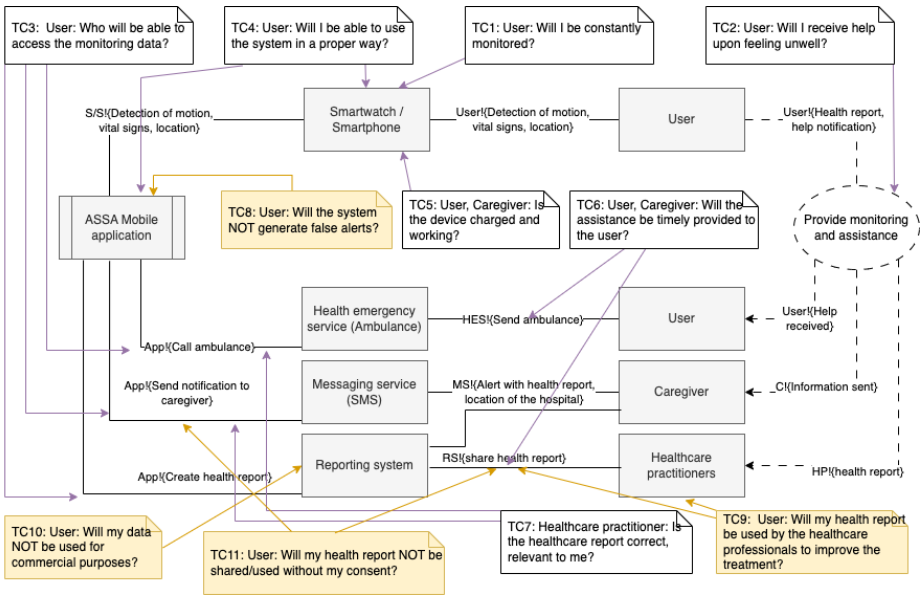

Fig. 3. Global system model of ASSA annotated with trust concerns from the Iteration 1 (TC1-TC7) and Iteration 2 (TC8-11).

Step 5: We refine the global system model by decomposing the main problem 'Provide monitoring and assistance' into the sub-problems corresponding to the

use-cases of ASSA defined in coordination with the project developers:
P1: Monitor the user and detect emergencies. This sub-problem consists of continuously collecting data about the user's vital parameters from the sensors (e.g., smartwatch, smartphone) and generating user's health metrics. Data are analyzed by the dedicated algorithms (i.e., compared to historical data or some baseline established by a physician). If some irregularities are detected, the alert is triggered.
P2: Send ambulance to the user if an emergency is detected Once the alert is triggered, the system transmits the alert to a healthcare emergency service, which sends an ambulance in response. Healthcare practitioners (e.g., paramedics, physicians) intervene to provide help to the user. User data related to the triggered alert (i.e., the health report) is communicated to healthcare professionals through the reporting system.
P3: Inform the caregiver about the emergency. The system notifies the caregiver via a text message (e.g., an SMS), providing her with the data related to the alert and the emergency intervention that followed (e.g., the address of the hospital where the user was transported).
P4: Manage and share health reports. The system records the user's health metrics creating regular health reports and incident health reports in an on-line Reporting system. With the user's authorization, these health reports can be communicated with the caregivers and the healthcare professionals for emergency interventions and regular check-ups.

We illustrate the refined system diagram for the sub-problem P2 in Fig. 4. TA1.01-04, TA1.08, TA1.09 are related to the ASSA Mobile system, smartwatch and smartphone – the elements that will be directly manipulated by the user. TA1.05-07, TA1.10 are linked to the external systems in the machine domain (i.e., Health emergency service, Messaging service, Reporting system). These are assumptions about how the ASSA Mobile application should be integrated with these systems and about specific properties or functionalities that these external systems must ensure. TA1.04, TA1.10-11 are related to the entities that are not included in the problem model – Network provider and Healthcare authorities.
Step 6: Using the TA, we formulate 11 trustworthiness requirements (1.01-1.11 in Table 3). We link the TwR with the TC they are addressing. Each TwR reflects one or more TA. For example, the requirement TwR1.05: The system shall be integrated with (recognized by) the health emergency service (e.g., Ambulance or SAMU in France) is associated with TA1.05: Emergency calls are responded 24/7 and TA1.10: The healthcare provider handles emergency alerts received from the system (Table2).
Step 7: The TwR are formulated for ASSA Mobile application, for the external systems and services (e.g., health emergency service, reporting system), and for the entities in the domain environment (e.g., healthcare practitioner, healthcare authorities, mobile network provider). The TwR related to ASSA mobile application can be considered as functional or non-functional requirements. Some of these requirements align with the As-Is design solution (indicated '+' in Table2), whereas the others lead to new design features. For example, TwR1.01, TwR1.03, TwR1.07 suggest the update in the ASSA mobile interface As-Is. The TwR related to external systems and services can be considered as non-functional, in-

tegration requirements, and regulatory and compliance requirements. They also introduce new elements to the design. For example, TwR1.09 focuses on the integration and compliance with the (existing) platforms for sharing medical data. TwR1.05-06, TwR1.08 highlight the importance of service-level agreements with healthcare and mobile network providers.

5.2 The Prospective Trust Analysis

The objective of this analysis is to extend our understanding of the system and its environment, focusing on the trustworthiness and acceptance of the system by its stakeholders. Compared to the retrospective analysis, here we apply the method to the newly collected empirical data for prospective TwR elicitation.
Step 1: We use the global system model of ASSA developed in the Iteration 1 (Fig. 2) and proceed with TC extraction.
Step 2: Through qualitative data analysis we were able to confirm the TC identified in the Iteration 1 and to identify new TC (see Table 1). In particular, the interviews reveal that the prospective users are concerned with a possibility of a false alert and the purposeful use of the collected data (see TC8-11, Table 1).
Step 3: We made the new trust assumptions about the (technical) properties of the system To-Be, which, if implemented, would alleviate the TC expressed by the users (TC8-11) and improve system trustworthiness. Note, that some trust concerns are addressed by the TA formulated in the Iteration 1. For example, TA1.03, 03.01 are already addressing the trust concern about the false alerts (TC8). The new TA are listed in Table 2 (see TA2.01-2.05). For example, we assume that the user can trust the system if she can learn to use the system from the documentation/support materials provided (TA2.07).
Step 4: We associate new TC and TA with the system components. Fig. 4 illustrates the traceability between the TA and the system elements for the system model for P2: Send ambulance to the user if an emergency is detected.
Step 5: We update the sub-problems (see the retrospective analysis) and their system models (omitted in this paper).
Step 6: We formulate five TwR (TwR2.01-2.05) and link them with their corresponding TA and TC.
Step 7: While TwR2.01 aligns with the ASSA As-Is design solution, the other four TwR are new. They have been validated by the ASSA developers and led to the design update. TwR2.02, TwR2.05 require the extension of the ASSA application interface whereas TwR2.03-04 need to be addressed by the external Reporting system and by the Healthcare practitioners (integration, regulatory and compliance requirements).

5.3 Discussion

Our study shows that the proposed method is applicable in retrospective (following up on a design process not focused on trust) and in prospective (being integrated in a method such as human-centered design), validating H1. We examined the identified TwR with the ASSA developers: five TwR correspond to the ASSA requirements As-Is; six new TwR led to the ASSA specification update

Table 3. Trustworthiness requirements

TwR	Description	Associated with TA	Addressing TC:	Included into ASSA?
TwR1.01	The system shall provide the user with the means to control the sensors status	TA 1.02	TC1	new / val
TwR1.02	The system and the watch shall have a simple and clear interface	TA1.08	TC4	+
TwR1.03	The system shall alert the user when the battery charge is low	TA1.09	TC5	new / val
TwR1.04	The system shall be compatible with certified/reliable sensors/components	TA 1.01/2.01	TC1, TC8	+
TwR1.05	The system shall be integrated with (recognized by) the health emergency service (e.g., Ambulance or SAMU in France)	TA1.10, TA1.05	TC2, TC6, TC9	new / val
TwR1.06	The system shall be certified / approved by the healthcare authority	TA1.11; TA1.03	TC2, TC7, TC8, TC9	new/val
TwR1.07	The user shall be able to trigger an alert manually	TA1.03.01	TC2, TC8	new/ inval
TwR1.08	The system shall use a reliable mobile/internet network provider	TA1.04; TA1.04.01	TC2	new/val
TwR1.09	The system shall use a report format compatible with cloud reporting systems	TA1.06-07 TA2.02-04	TC3, TC9-11	new / val
TwR1.10	The system shall encrypt all data, stored and exchanged	TA1.06	TC3	+
TwR1.11	The system shall be integrated with an on-line Reporting system	TA1.07	TC3, TC10-11	+
TwR2.01	The system has to be GDPR-compliant	TA2.02	TC10-11	+
TwR2.02	The user shall be able to control report sharing	TA2.02; TA1.07	TC3, TC10-11	new/val
TwR2.03	The reporting system must be highly available	TA2.04-05	TC9	new/val
TwR2.04	The healthcare practitioner has to use the received health report while taking the user in charge	TA2.05; TA2.06	TC9	new/val
TwR2.05	The user shall be able to access training materials and guidelines for the ASSA mobile application online (e.g., FAQ, getting started videos etc.)	TA1.02, TA1.03.01, TA1.08-09	TC1-2, TC4-5, TC8	new / val

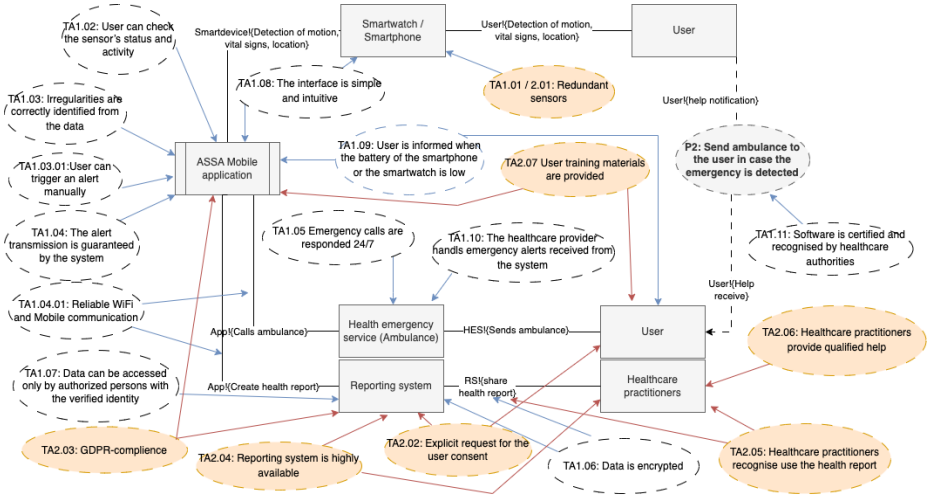

Fig. 4. Refined system diagram for the sub-problem P2: Send ambulance to the user in case the emergency is detected. Trust assumptions from the Iteration 1 (TA1.xx) and from the Iteration 2 (TA2.xx) are linked to the system elements.

(Table 3). This validates our research hypothesis H2, H3 and demonstrates the method efficiency.

We formulated 18 trust assumptions based on the retrospective and prospective trust analysis (Table 2) and created a traceability matrix and system models explicitly linking the (social) trust concerns and the system elements. We conducted feedback sessions with the engineers to ensure a shared understanding and an added value of this traceability, validating H4.

We formulated 11 new TwR with 10 recognized important and validated by the ASSA developers. TwR 1.07: 'The user shall be able to trigger an alert manually' was not validated as it goes against the product vision of ASSA, where the system takes the whole responsibility for the emergency detection. By this, we were able to partly validate H5 and the method relevance.

6 Conclusions

In this work, we applied the method for iterative trustworthiness requirements analysis and elicitation elaborated from [16] to the case study of the ASSA CPS. We formulated trust assumptions that justify technical decisions, supporting alignment between (social) trust concerns and specific properties of the system, and addressing the world-on-machine paradox formulated in [26]. The conducted trust analysis led to updates in the ASSA solution design, with a focus on enhancing system trustworthiness. The ASSA developers acknowledged the significance of this analysis.

In future work, we intend to evaluate the proposed design to assess the impact of TwRs on the perceived trustworthiness of the system. To achieve this, we plan to conduct user-centered evaluations and post-implementation usability testing.

References

1. Abtoy, A., Touhafi, A., Tahiri, A., et al.: Ambient assisted living system's models and architectures: A survey of the state of the art. Journal of King Saud University-Computer and Information Sciences **32**(1), 1–10 (2020)
2. Amaral, G., Guizzardi, R., Guizzardi, G., Mylopoulos, J.: Ontology-based modeling and analysis of trustworthiness requirements: Preliminary results. In: International Conference on Conceptual Modeling. pp. 342–352. Springer (2020)
3. Amaral, G., Guizzardi, R., Guizzardi, G., Mylopoulos, J.: Trustworthiness requirements: the pix case study. In: Conceptual Modeling: 40th International Conference, ER 2021, Virtual Event, October 18–21, 2021, Proceedings 40. pp. 257–267. Springer (2021)
4. Caballero, P., Ortiz, G., Medina-Bulo, I.: Systematic literature review of ambient assisted living systems supported by the internet of things. Universal Access in the Information Society pp. 1–26 (2023)
5. Gambetta, D., et al.: Can we trust trust. Trust: Making and breaking cooperative relations **13**(2000), 213–237 (2000)
6. Garry, T., Harwood, T.: Trust and its predictors within a cyber-physical system context. Journal of Services Marketing **33**(4), 407–428 (2019)
7. Haley, C.B., Laney, R.C., Moffett, J.D., Nuseibeh, B.: The effect of trust assumptions on the elaboration of security requirements. In: Proceedings. 12th IEEE International Requirements Engineering Conference, 2004. pp. 102–111. IEEE (2004)
8. Haley, C.B., Laney, R.C., Moffett, J.D., Nuseibeh, B.: Picking battles: The impact of trust assumptions on the elaboration of security requirements. In: Trust Management: Second International Conference, iTrust 2004, Oxford, UK, March 29-April 1, 2004. Proceedings 2. pp. 347–354. Springer (2004)
9. Haley, C.B., Laney, R.C., Moffett, J.D., Nuseibeh, B.: Using trust assumptions with security requirements. Requirements Engineering **11**, 138–151 (2006)
10. ISO/IEC: 9241-210:2019(en) ergonomics of human-system interaction — part 210: Human-centred design for interactive systems. Tech. rep. (2019)
11. Kitchenham, B., Pickard, L., Pfleeger, S.L.: Case studies for method and tool evaluation. IEEE Softw. **12**(4), 52–62 (1995)
12. Lee, E.A.: Cyber physical systems: Design challenges. In: 2008 11th IEEE international symposium on object and component-oriented real-time distributed computing (ISORC). pp. 363–369. IEEE (2008)
13. Mayer, R.C., Davis, J.H., Schoorman, F.D.: An integrative model of organizational trust. Academy of management review **20**(3), 709–734 (1995)
14. Mcknight, D.H., Carter, M., Thatcher, J.B., Clay, P.F.: Trust in a specific technology: An investigation of its components and measures. ACM Transactions on management information systems (TMIS) **2**(2), 1–25 (2011)
15. Meeßen, S.M., Thielsch, M.T., Hertel, G.: Trust in management information systems (mis). Zeitschrift für Arbeits-und Organisationspsychologie A&O (2019)
16. Mohammadi, N.G.: Trustworthy cyber-physical systems: A systematic framework towards design and evaluation of trust and trustworthiness. Springer Vieweg, Wiesbaden, Germany, 1 edn. (2019)
17. Mohammadi, N.G., Ulfat-Bunyadi, N., Heisel, M.: Problem-based derivation of trustworthiness requirements from users' trust concerns. In: 2018 16th Annual Conference on Privacy, Security and Trust (PST). IEEE (2018)
18. Neumann, P.G.: Fundamental trustworthiness principles. New Solutions for Cybersecurity (2018)
19. Reichstein, C., Härting, R.C., Häfner, F.: Challenges in the market launch of active assisted living solutions-empirical results from european experts. Procedia Computer Science **176**, 2000–2009 (2020)

20. Rousseau, D.M., Sitkin, S.B., Burt, R.S., Camerer, C.: Not so different after all: A cross-discipline view of trust. Academy of management review **23**(3), 393–404 (1998)
21. Rychkova, I., Ghriba, M.: Trustworthiness requirements in information systems design: Lessons learned from the blockchain community. Complex Systems Informatics and Modeling Quarterly (35), 67–91 (2023)
22. Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., Leimeister, J.M.: Understanding the formation of trust in it artifacts (2012)
23. Stokke, R.: The personal emergency response system as a technology innovation in primary health care services: an integrative review. Journal of medical Internet research **18**(7), e187 (2016)
24. Sutcliffe, A., Sawyer, P., Bencomo, N.: The implications of 'soft'requirements. In: 2022 IEEE 30th International Requirements Engineering Conference (RE). pp. 178–188. IEEE (2022)
25. Ulfat-Bunyadi, N., Meis, R., Heisel, M.: The six-variable model-context modelling enabling systematic reuse of control software. In: International Conference on Software Paradigm Trends. vol. 2, pp. 15–26. SCITEPRESS (2016)
26. Wang, X., Mylopoulos, J., Guizzardi, G., Guarino, N.: How software changes the world: The role of assumptions. In: 2016 IEEE Tenth International Conference on Research Challenges in Information Science (RCIS). pp. 1–12. IEEE (2016)
27. Warwick, L.: Designing trust: the importance of relationships in social contexts. The Design Journal **20**(sup1), S3096–S3105 (2017)