On The Quest of Trust Requirements for Socially Assistive Robots

Larissa Costa
Centro de Informática
Universidade Federal de Pernambuco
Recife, Pernambuco, Brazil
lrc@cin.ufpe.br

Jaelson Castro
Centro de Informática
Universidade Federal de Pernambuco
Recife, Pernambuco, Brazil
jbc@cin.ufpe.br

Judith Kelner
Centro de Informática
Universidade Federal de Pernambuco
Recife, Pernambuco, Brazil
jk@cin.ufpe.br

Bruno Jeronimo
Centro de Informática
Universidade Federal de Pernambuco
Recife, Pernambuco, Brazil
bsj@cin.ufpe.br

Maria Lencastre
Escola Politécnica de Pernambuco
Universidade de Pernambuco
Recife, Pernambuco, Brazil
mlpm@ecomp.poli.br

Óscar Pastor
Universitat Politècnica de València
Valencia, Spain
opastor@dsic.upv.es

***Abstract*—The development and use of Socially Assistive Robots has grown significantly in recent years. Trust can be defined as the user's belief that the robot will fulfil its expected behaviour in a predictable, effective, and secure manner. Hence, trust is one of the critical aspects for their adoption in a social setting. Our goal in this research is to develop a catalogue of trust requirement of Anthropomorphic Socially Assistive Robots. In this paper we describe the process used for the catalogue construction, which resulted in the identification and modelling of 125 requirements (14 main property) in NFR language and Specification Cards. Preliminary validation included its use for the development of Proof-of-Concept application which explores how the NAO Robot could assist in upper limbs motor rehabilitation as well as interview with 8 experts and an assessment with 20 Requirements Engineers. Early results indicate that catalogue effectively supported the elicitation and specification phases of socially assistive anthropomorphic robot projects, facilitating the identification of potential issues from the perspective of Trust. However, several points of improvements were also identified.**

***Keywords*— Requirements, Trust, Human-Robot Interaction, Socially Assistive Robot**

I. Introduction

Recently, the interest in assistive robotics has grown significantly [1]. Socially Assistive Robots (SARs) have demonstrated their effectiveness in various application domains, including therapy for neurodivergent children (such as those on the autistic spectrum) and as a tool in caring for the elderly or individuals with physical, cognitive, or social disorders. They have also shown promise in physical rehabilitation [2,3], among other areas. The field of Human-Robot Interaction (HRI) is responsible for investigating collaborative events between humans and robots, underlying the roles of such devices in our lives [4]. The nature of close interactions and collaboration between humans and socially assistive robots underscores the importance of investigating the requirements for establishing trust from the user's perspective [5].

Trust can be defined as the user's belief that the robot will fulfil its expected behaviour in a predictable, effective, and secure manner. Trust involves evaluating the risks associated with interacting with these robotic agents [6,7,8]. Consequently, it is crucial to consider the requirements associated with the social and psychological influences introduced by the robot during the interaction, as well as the spatial and physical contact variables in HRI scenarios.

The main objective of this research is to investigate the factors that contribute to human trust in robotic devices [9], such as SARs [10,11]. Our intention is to develop a comprehensive catalogue of Non-Functional Requirements (NFRs) for Trust tailored for Anthropomorphic-type SARs. This catalogue will assist robot developers in incorporating trust-related NFRs into their designs, help stakeholders in selecting suitable SARs based on trust requirements, and guide application developers in integrating SARs with the necessary trust considerations. To achieve this objective, we aim to answer the following research questions:

1) What are the main SARs Trust NFRs that need to be considered?

We hope to identify the essential trust-related requirements that need to be considered during the design and implementation of Anthropomorphic-type SARs. It will also provide guidance for those stakeholders that may need to select (e.g. to purchase or use) SARs as well as application developers for SARs, enabling them to focus on critical factors that contribute to establishing trust in these robotic systems.

2) Is the NFR Framework appopriate for modeling trust-related requirements in the context of Human-Robot Interaction?

Given that NFR Framework [12] is a popular notation for non-functional requirements modelling, we would like to examine if it is appropriate for representing trust requirements in the context of human-robot interaction scenarios.

This study is based on a literature review that explores the state-of-the-art in Socially Assistive Robots, Human-Robot Interaction, and Trust.

Preliminary validation included its use for the development of Proof-of-Concept (PoC) application which explores how the NAO robot could assist in upper limbs motor rehabilitation. We also conducted a limited number of interviews (8) with experts in the fields of Human-Robot Interaction, Trust, and Requirements Engineering. These interviews allowed for in-depth discussions and insights from experienced professionals in these areas, providing valuable feedback on the Trust Catalogue. In addition to the interviews with 8 experts, a survey was administered to hear the opinions of (20) Requirement Engineering (RE) students. This allowed for a broader perspective, gathering input from individuals who are studying and learning about the field. Early results indicate that catalogue effectively supported the elicitation and specification phases of socially assistive anthropomorphic robot projects, facilitating the identification of potential issues from the perspective of trust. However, several points of improvements were also identified.

The rest of this paper is organized as follows: Section 2 provides a summary of background information, Sections 3 and 4 explain the process used to develop the catalogue, Section 5 presents the catalogue itself, while Section 6 discusses its preliminary. Section 7 explores related works, Section 8 addresses potential threats to the validity of the research, and finally, Section 9 summarizes the findings and highlights opportunities for future work.

II. BACKGROUND

Human-Robot Interaction is a field dedicated to studying the interaction between humans and robots, with goals encompassing the design, comprehension, and evaluation of robotic systems intended for use by or with humans [13]. Within the HRI community, research efforts are driven by the quest to ensure safety in human-robot collaborations and understand Trust's role in such interactions [5,14,15].

While safety and trust are often treated as separate investigation aspects, they are interconnected. The perceived level of safety and security directly influences the level of trust individuals place in robot usage. Lower perceived safety and security lead to decreased Trust. Considering this correlation, we view safety as a subtopic of trust. In the literature, privacy-related aspects have received significant consideration regarding security in the user perspective of trust.

Trust is a crucial element in various disciplines, impacting the efficacy of communication, learning, and problem-solving [16]. However, defining trust is challenging, as its meaning can vary contextually, shaped by tacit experiences, social exchanges, and cultural nuances [16]. In our work, Trust is defined as the user's belief (trustor) that the robot (trustee) will fulfil its expected functions in a predictable, effective, and secure manner.

Socially Assistive Robotics (SAR) is a prominent domain within HRI, lying at the intersection of Assistive Robotics (AR) and Socially Interactive Robotics (SIR). SAR's primary objective is to facilitate proximate and effective interactions that enable users to achieve measurable progress in physical recovery, rehabilitation, learning, and other tasks.

Non-Functional Requirements, often considered as quality attributes, directly relate to system features. These requirements play a significant role in the development of robotic systems, directly impacting the outcomes of development efforts. Overlooking NFRs can lead to diminished system usability and effectiveness [12, 17, 18].

III. CATALOGUE CONSTRUCTION PROCESS

The catalogue was developed using the extracted knowledge to include essential trust-related non-functional requirements specifically tailored for Anthropomorphic-type SARs. The construction process of the catalogue involves three main phases: **Initial Search**, **Development**, and **Evaluation.** For further details, please refer to [25].

A. *Initial Search phase*

In the Initial Search phase, we conduct a non-systematic literature review to identify trust factors specifically related to SARs. We consulted various digital libraries. Keywords related to Socially Assistive Robots, Human-Robot Interaction, and Trust were used to search for most cited and relevant works. The snowballing technique was employed to identify additional papers through references and citations.

This involves:

1. Identification of relevant Taxonomies: In this task, we identified the relevant trust taxonomies in the HRI field. These taxonomies served as a basis to localize the trust factors pertinent to SARs.
2. Identification of Trust Factors for SARs: Through a research effort, we identify the key trust factors that are relevant to SARs.
3. Taxonomy Adjustment: Existing trust taxonomies are reviewed and adapted to align with the characteristics of SARs.
4. Identification of NFRs according to the main properties of SARs: We map the identified trust factors to the main properties of SARs to determine the NFRs associated with trust.
5. Collection of Definitions and Attributes of NFRs: We gather definitions and attributes of the identified NFRs from various sources, including research papers, books, monographs, dissertations, and theses.

B. *Development phase*

The Development phase focussed on organizing the NFRs and creating the catalogue:

6. Organization of the NFRs: We make necessary modifications to the nomenclature and categorize the NFRs based on the identified properties of SARs.
7. Creation of the Catalogue: The filtered NFRs are synthesized and compiled into a comprehensive catalogue. We used the NFR framework notation, such as the Softgoal Interdependence Graph (SIG), to visually depict the interdependencies among the NFRs.

C. *Evaluation phase*

The Evaluation phase ensures the quality and effectiveness of the constructed catalogue:

8. Proof of Concept Application: The catalogue was applied in a collaborative study with the author of [19] that employed a Socially Assistive Robot (NAO) for upper limb motor rehabilitation.
9. Evaluation of the PoC and the Catalogue with Specialists: Interviews were conducted with experts

in the field of SARs, Human-Robot Interaction, and the NFR framework to evaluate both the PoC and the catalogue.

10. Further Evaluation of the Catalogue with Requirements Engineers Students: To complement the evaluation process, an online questionnaire was administered to a group of students specializing in Requirements Engineering to assess the catalogue. The students were given access to the catalogue and asked to provide feedback on its structure, clarity, and overall usefulness.

IV. RESULTS OF THE CATALOGUE CONSTRUCTION STAGES

During the initial search, we discovered three relevant taxonomies in the field of Trust-Human-Robot Interaction: [20,21,22]. In addition, we came across the work of [23], which focuses on trust factors in Socially Assistive Robots for rehabilitation. While Langer et al. [23] provide design guidelines and measurement methods for trust, they did not propose a new taxonomy.

To further enrich our understanding of trust factors specific to SARs, we integrated the insights from Langer et al. [23] with the existing taxonomies. This comprehensive approach enabled us to adapt the taxonomy for SARs and identify 30 relevant trust factors (see in the adapted taxonomy[1]). These factors were carefully examined and investigated within the scope of the existing taxonomies, contributing to a more comprehensive understanding of trust in SARs.

The taxonomy, as presented in the provided link[1], is structured into three general categories: **Human factors**, **Robotic factors**, and **Environmental factors**, further subdivided into eight sub-categories. These categories provide a comprehensive framework for understanding the various aspects influencing trust in Socially Assistive Robots.

During Task 4, we organized the information and established specific objectives to handle the many factors in the trust taxonomy. As a result, we identified twelve SARs Properties [24] that specifically highlight the unique characteristics of Socially Assistive Robots. These properties were a solid foundation for our subsequent analysis and categorization of trust factors.

In addition to the SARs properties, we recognized the utmost importance of safety and privacy in the design and implementation of Socially Assistive Robots, which has been emphasized in studies such as [23]. Our catalogue incorporates primary non-functional requirements dedicated explicitly to Privacy and Safety. These NFRs have been thoughtfully included based on their significant role and serve as guidance for engineers and designers, ensuring the development of trustworthy and responsible robotic systems.

Furthermore, as we delved into the trust taxonomy, we associated 19 out of the initial 30 factors with the SARs properties [24]. Each factor was meticulously aligned with its corresponding definitions, leading to a comprehensive understanding of the trust dynamics in the context of Socially Assistive Robots.

During Task 5, we aimed to gather the definitions and attributes of the non-functional requirements. Subsequently, we moved on to the Development phase, which consisted of: Organizing the identified NFRs (Task 6) and Creating the catalogue (Task 7). Task 6 involved making nomenclature modifications and organizing the factors according to the SARs properties. As a result, we established the primary NFRs as follows: **Appearance/Anthropomorphism, Communication, Emotion, Sophistication of Interaction, Socially situated learning, Human-oriented perception, Privacy, Safety, Personality, Intentionality, Role of the Robot, Tasks/Context, User modelling and User Populations** (see Fig. 1).

In Task 7, we constructed the Catalogue of Non-Functional Requirements by organizing and documenting all the filtered NFRs from the previous tasks. We carefully considered the definitions, attributes, and restrictions of each NFR. While all the factors addressed in the taxonomy are important in building trust, this catalogue primarily focuses on the factors related to the robot. To enhance the structure of the catalogue, we adopted the NFR framework notation proposed by Chung, Nixon and Mylopoulos (2000) [12], which is widely used for representing catalogues in Requirements Engineering. We created a Softgoal Interdependence Graph that visually represents the interdependencies among the non-functional requirements, providing a comprehensive overview of their relationships. For further details please refer to [25].

V. THE CATALOGUE

The NFR4TRUST (Non-Functional Requirements for Trust) catalogue contemplates the following supporting artifacts: a Softgoal Interdependency Graph that underlines the hierarchical relationship of requirements; a Correlation Table that links the primary requirements with the definitions for each NFR; and a set of Specification Cards. Fig. 1 presents a partial vision of the generated SIG, emphasizing the 14 primary NFRs without the refinements and correlations. However, if we include their refinements, a total of 125 requirements were identified [25]. Note that the nomenclature used in the NFR4TRUST catalogue is aligned with the terminology commonly used in the field of Human-Robot Interaction literature. By utilizing consistent terminology, we ensure that the catalogue is in line with established practices and conventions within the HRI community. The complete version of the SIG and tables indicating the NFRs correlation is available at the Link[1].

Due to space limitations, we can only present some of the 125 requirements. Hence, we will provide specification cards overview of the 14 primary requirements. Among them, we will focus on the *Appearance/Anthropomorphism* requirement.

[1] The taxonomy is available at the link: https://drive.google.com/file/d/1X3MZqLlqjCpgKSDVepyy1auhvRMys17o

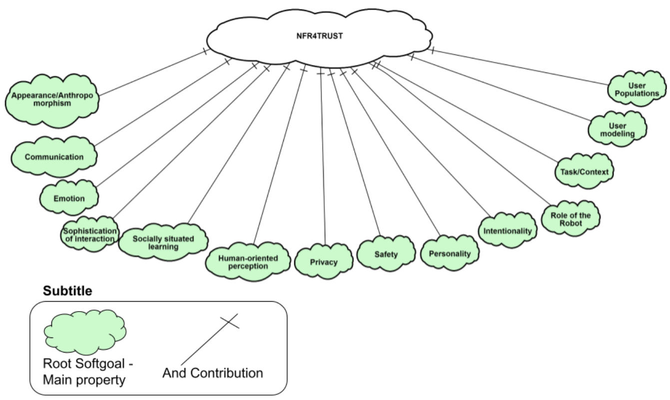

Fig. 1. Partial view of the SIG of the primary NFRs

A. *Appearance/Anthropomorphism*

As an example, in Fig. 2, we provide the refinement of the primary Appearance/Anthropomorphism NFR.

Definition: The shape and structure of the robotic device, both elements that constitute its appearance, are essential to establish social expectations [26]. According to Duffy (2003) [27], anthropomorphism is the tendency to attribute human characteristics to inanimate objects, animals, and other artifacts. In this same fashion, the concept of anthropomorphism plays a relevant role in HRI. In the SARs project, the robot's general appearance may be subject to change to support the tasks and context of use. Systems designers may also consider environmental variables and user population as long as the physical attributes fulfil the user's expectations [27]. Consequently, the robot's physical features and anthropomorphism level are paramount to supporting trust, particularly when such characteristics enhance social perceptions.

Main attributes: Functional capabilities, physical characteristics. Attributes are also defined, see [25].

Literature sources: [21, 26, 27, 28, 29]

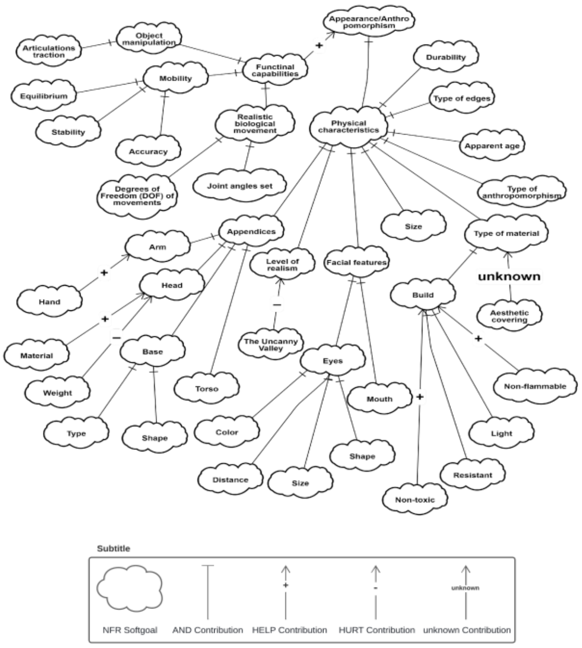

Fig. 2. Refinement of the Appearance/ Anthropomorphism NFR.

Next, the two main Appearance/Anthropomorphism attributes are defined. All its sub-attributes are described in [25].

Functional Capabilities

Definition: It refers to the robot's ability to perform actions and tasks despite its technological and design limitations.

Literature sources: [21,26,27,30]

Sub-attributes: Object manipulation, Mobility, Realistic biological movement.

Physical characteristics

Definition: They are the external qualities of a being. The characteristics should be aligned with the functional capability, context, and task, in addition should avoid creating exaggerated expectations in the user.

Literature sources: [31,32]

Sub-attributes: Appendices, Level of realism, Facial features, Size, Type of material, Type of anthropomorphism, Apparent age, Type of edges, Durability.

Correlations: In Figure 3, we can see the general correlations (dotted line) of Appearance/Anthropomorphism. We presented four "unknown" correlations among *Appearance/Anthropomorphism* and *Emotion*, *Safety*, *Role of the robot*, and *User populations* requirements. We also found three **positive** correlations towards Personality, Human-oriented perception, and Intentionality requirements. These correlations represent the impacts that Appearance/Anthropomorphism generates on these requirements.

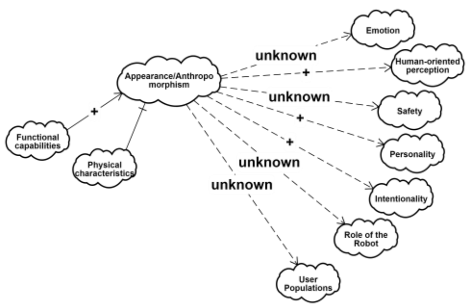

Fig. 3. Appearance/Anthropomorphism NFR- contributions and correlations.

Next, we briefly describe the remaining 14 primary NFRs depicted in Figure 1. The correlations between these NFRs can be found in the table included in the supporting materials [25].

B. *Communication*

Definition: The robot's communication is an essential process in the interaction, involving exchanging information from the robot to the user, enabling the creation of messages that elicit a response. Communication regarding accuracy, feedback, cues, modes, and access to information is essential for building trust.

Main attributes: Feedback, Mode of communication, Reliability, Accuracy, Proximity, Turn-Taking.

C. Emotion

Definition: Emotions play a crucial role in human interactions. In human-robot interactions, emotions help facilitate a more believable and engaging interaction.

Main attributes: Emotional Postures, Facial Expressions, Emotionally Intelligent.

D. Sophistication of Interaction

Definition: This requirement focuses on the user's interaction with the robot and the robot's ability to process information from the reciprocal interaction. The level of sophistication in the interaction directly impacts the robot's social perception. However, achieving a higher level of sophistication may require increased technical capacity.

Main attributes: Interaction modalities.

E. Socially situated learning

Definition: This requirement pertains to the robot's capability to acquire new skills through social interaction with the environment or the user. Socially situated learning enables the robot to learn and adapt its behavior based on the social context.

Main attributes: Robot social learning.

F. Human-oriented perception

Definition: This requirement focuses on the robot's ability to perceive and process information from the environment, particularly from the user, in a manner that closely resembles human perception. It involves the robot's capacity to interpret sensory inputs and understand human cues and signals.

Main attributes: Automation.

G. Privacy

Definition: It concerns the requirements related to the protection of personal information of the user and his environment against access by unauthorized persons.

Main attributes: Confidentiality.

H. Safety

Definition: This requirement focuses on ensuring the physical safety of both the user and the robot itself. It involves measures to prevent harm, injury, or damage during the interaction between the robot and the user.

Main attributes: Physical damages, Self-protection.

I. Personality

Definition: This requirement focuses on the robot's ability to demonstrate characteristics that resemble human personality traits. It involves the robot's behavior, expressions, and responses that exhibit emotions, attitudes, and dispositions similar to those of humans. The robot's personality enhances its ability to engage and establish rapport with users, making the interaction more natural and enjoyable.

Main attributes: Adaptability

J. Intentionality

Definition: This requirement pertains to the robot's ability to demonstrate behaviors that indicate intention in its actions. Intentionality refers to the robot's capability to exhibit purposeful and goal-directed behaviors, showing that its actions are driven by internal states and objectives. This attribute enhances the robot's perceived agency and promotes a sense of predictability and reliability in its interactions with users.

Main attributes: Robot Behaviour.

K. Role of the robot

Definition: This requirement defines the function performed by the robot in relation to its interaction with the user. It specifies the role that the robot assumes in the user's context, such as being an assistant, tutor, partner, or any other designated role. The robot's role influences its behavior, capabilities, and the nature of its interactions with the user.

Main attributes: Authority.

L. Task/Context

Definition: This requirement concerns the context of the use of the robot and the specific activities that the robot will carry out. It involves understanding the tasks the robot will undertake, the environment in which it will operate, and the context in which the interactions with the user will occur. The tasks and context significantly influence the design and implementation of the robot's functionalities and behaviors.

Main attributes: Mitigating Risks/Uncertainties, Establishing Context/Task Type, Physical Environment constraints, Setting Task complexity.

M. User modeling

Definition: This requirement specifies the set of user characteristics that will be served by the robot, considering demographic features, special needs, and relevant factors influencing interaction. The robot should be designed and tailored to accommodate these user populations' diverse needs and preferences, ensuring inclusivity and providing a personalized and effective user experience.

Main attributes: Understanding ability to use, Understanding prior experience.

N. User Populations

Definition: The robot's communication is an essential process in the interaction, involving exchanging information from the robot to the user, enabling the creation of messages that elicit a response. Communication regarding accuracy, feedback, cues, modes, and access to information is essential for building trust.

Main attributes: Understanding special needs, Establishing age group.

In Section 6 of our research, we present the results of the third phase of our work, which focused on evaluation. We conducted various evaluation activities to gather feedback and insights from experts and stakeholders in the field of HRI, RE and Trust.

VI. THE CATALOGUE ANALYSIS

This section presents the pilot evaluation of the catalogue. Firstly, we begin by characterizing the use of the catalogue in developing SARs applied in upper limb rehabilitation via proof of concept. Then, we report the proposed catalogue and PoC evaluation with the experts and specialists. Additionally, we will describe the results obtained from a questionnaire administered to students specializing in requirements engineering.

A. *Evaluation via Proof-of-Concept*

It is known that robots with social characteristics allow patients to create affective bonds, increasing the motivation to continue the treatment, making it an excellent strategy for long-term tasks present in physical therapy. We chose the NAO V5 robot as our anthropomorphic socially assistive robot for the proof of concept. We explored how it could assist in upper limbs motor rehabilitation. The catalogue was used to build the NAO Physio application [19].

Fig. 4 illustrates part of the Appearance/Anthropomorphism SIG instantiation. In this excerpt, we highlight the presence of operationalization (dark clouds, represent implementation/solutions), statements (dotted clouds, represent rationale), and contributions (positive, negative, AND must be satisfied). The primary Appearance/Anthropomorphism NFR has two attributes: Functional capabilities and Physical characteristics. In the Functional capabilities attribute, we have a positive contribution to the Appearance/Anthropomorphism Softgoal. The Physical characteristics attribute has an AND contribution to Appearance/Anthropomorphism. The SIG containing the instantiations related to Appearance/Anthropomorphism in the proof of concept can be accessed at Link[1]:

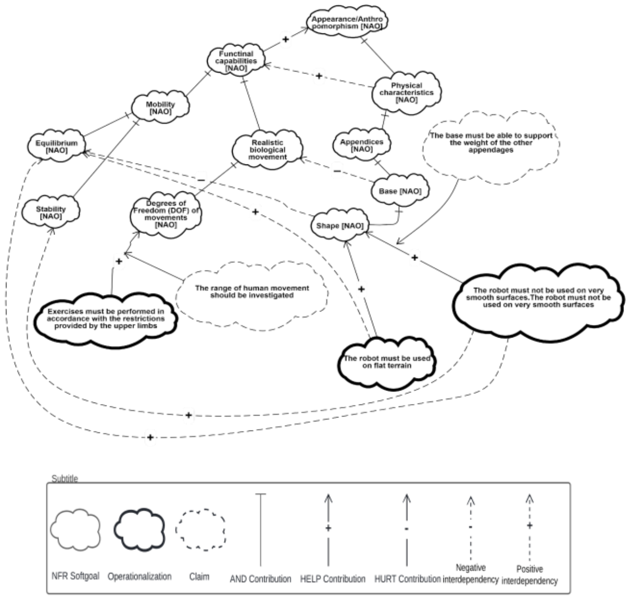

Fig. 4. Part of the Appearance/Anthropomorphism SIG instantiated in the Proof of Concept.

Overall, the **NAO Physio** [2] PoC provided an opportunity to evaluate the practical utility of the catalogue in decision-making processes related to robot selection and application development, with a specific focus on trust-related considerations.

B. *Preliminary Evaluation with RE and HRI Experts*

We conducted remote interviews with two senior experts with more than ten years of experience in their respective fields: a Requirements Engineer with NFR Framework knowledge and a HRI Engineer with Social Robots knowledge. These interviews were conducted via videoconference, with each session lasting one hour and thirty minutes on average. These interviews were recorded for future reference and analysis.

On one hand, the evaluation with RE expert was to ensure the appropriateness of the applicability of the catalogue in real scenarios as well as to check the correctness the proposed SIG. This evaluation made it possible to demonstrate the relevance of the NFRs for the domain of SARs and the feasibility of applying the catalogue in real scenarios. Overall, the Requirements Engineering expert considered that the use the NFR Framework was appropriate, including the correct application of the operationalization concept.

On the other hand, the appraisal with HRI expert emphasized the catalogue's usefulness for the design of SARs applications and its adoption for the developer of applications in this field. Privacy was identified as the most crucial factor in Socially Assistive Robots. provided suggestions for improvement. Further details can be found in previous works [25].

Some suggestions for improvement were made by the RE and HRI experts. Most were incorporated to the catalogue.

C. *Catalog evaluation with Further Specialists*

Once the catalogue was revised, we conducted a new round of interviews involving three additional experts from Requirements Engineering and three from Human-Robot Interaction.

We conducted semi-structured interviews based on the guidelines in [33]. These interviews provided a flexible yet guided approach, allowing for both structured questions and conversational exploration. The interviews were conducted remotely on videoconference, tailored to the expertise of each participant, and lasted between one and two hours. We recorded the interviews for future reference and analysis.

The interviewed specialists were primarily from Brazil and Canada, with significant experience and advanced qualifications. When assessing their knowledge outside their primary domain, we found that both RE and HRI specialists had limited knowledge of the other area. HRI specialists demonstrated a moderate level of knowledge in Requirements Engineering, while RE specialists had a lesser understanding of Human-Robot Interaction.

The RE specialists showed a high level of knowledge in requirements elicitation and specification, with varying familiarity with the NFR Framework. Surprisingly, their knowledge of social robots was moderate, while their understanding of Trust (in HRI) was limited. However, all participants recognized the significance of Trust for the acceptance of Social Robots. In contrast, the HRI specialists demonstrated a strong understanding of Social Robots but had more moderate knowledge of Socially Assistive Robots. Their self-assessment of Trust knowledge showed varied responses, but most considered it to be highly important for Social Robot acceptance.

After assessing the participants' profiles, we provided a concise overview of the research topics, focusing on areas where their knowledge was lacking. For the RE specialists,

[2] The name of the NAO robot aimed at physiotherapy

we presented the catalogue, including correlations and an example of the proof of concept (Appearance/Anthropomorphism).

The RE experts evaluated the catalogue by responding to 12 questions, including 5 Likert-scale questions. The HRI specialists were presented with 8 questions regarding the catalogue, with only 3 using the Likert scale.

D. *Qualitative analysis*

During the qualitative analysis, we transcribed the specialist's opinions and created codes based on their responses: "Agree" when they fully agreed with the questions, "Agree (but with reservations)" when they agreed but had some caveats, and "Disagree" when they did not agree at all.

Both RE and HRI experts provided several improvement considerations. They suggested verifying the possible evolution of the NFR Framework language to cover questions related to robot interaction. Additionally, they proposed mapping requirements associated with robot awareness, conversational agents, and psychological factors in conversations. To manage the complexity of the catalogue, the experts recommended organizing it with different levels of abstraction or views.

They also suggested presenting properties in more textual forms or selectable aspects based on relevance. Abstracting decompositions was another suggestion put forward. The specialists emphasized the importance of incorporating empathy and addressing robot gender and cultural issues in the catalogue. It was also emphasized the need to use appropriate terminology as well as the necessity to better organize/structure the information to improve its usability for non-experts.

E. *Expert´s assessment of taxonomy human-robot interaction*

Although it is not the focus of this research, we have evaluated the factors and organization of the Trust taxonomy for SARs used as the basis of our catalogue. This assessment was carried out only with HRI experts and aimed to assess the importance, impacts and priorities of the factors, and the organization and completeness of the taxonomy information. This assessment was limited in scope, as it involved only a small group of HRI experts. As such, the results should be interpreted cautiously, and further investigation with a more extensive and diverse sample would be beneficial.

Table III presents the 6 questions related to the taxonomy and the respective answers of the HRI specialists. It is challenging to draw definitive conclusions from these results. The interpretation of the factors and their priorities may vary depending on the specific context and use case envisioned by each expert. Therefore, it is essential to consider the individual perspectives and expertise of the specialists when interpreting the findings.

TABLE III. Answers from Human-Robot Interaction experts on taxonomy-related questions.

QUESTION	EXPERT HRI_1	EXPERT HRI_2	EXPERT HRI_3
Q1	Human Factor	Robot Factor	Cultural/Societal Impact
Q2	Environment	Expectancy	Environment
Q3	Agree(but with reservations)	Agree(but with reservations)	Agree(but with reservations)
Q4	Appearance/Anthropomorphism	Appearance/Anthropomorphism	Human Factor
Q5	Time, Laws in general, Compliance; Agency/Agents model	Ethics, Transparency, Clarity, Expressive communication, Quality of interaction	Social acceptance
Q6	The form of Taxonomy representation; Put the legal framework on top of trust.	Break communication into a few more levels; Improve the robot factor Division.	Break down the factors into more levels.

Q1 - Among the trust factors, which do you think is the most important?
Q2 - Among the trust factors, which one do you think is the most difficult to achieve?
Q3 - Do you agree with the organization of the taxonomy?
Q4 - Which factor present in the taxonomy do you think would most negatively impact user trust if not taken into consideration?
Q5 - Are there any factors you would like to add to the taxonomy?
Q6 - Would you change anything in the taxonomy?

F. *Discussion of Results*

Before delving into the discussion of the obtained evaluation results, it is crucial to highlight that the validation process of the NFR4TRUST Catalogue primarily focused on capturing the subjective "perception" of utility as reported by experts and students. This emphasis was placed on gathering qualitative feedback and insights rather than measuring its utility through objective metrics or quantitative assessments.

One important factor to consider is the influence of the specialists' experience level on their responses. Highly experienced RE specialists tended to exhibit a more neutral or reluctant stance toward the evaluations. Their strict adherence to the rules of their field made it challenging to develop multidisciplinary content that accommodated everyone's understanding. It remains uncertain whether the participants' extensive experience influences this characteristic.

Additionally, the limited knowledge of specialists in areas outside their expertise posed difficulties in evaluating the catalogue comprehensively. Understanding essential concepts, such as specific robotics requirements for RE experts and the notation rules of the NFR Framework for HRI experts, proved challenging.

Considering the experts´ backgrounds, it is important to acknowledge the possibility of the Dunning-Kruger effect [26], where participants may overestimate their competencies. This effect could explain the overly positive responses from some participants.

Regarding the participants' knowledge of the NFR Framework, despite all RE experts considering their knowledge level to be very high, discrepancies in feedback on the notation rules were observed. Whether this is due to the Dunning-Kruger effect or extensions in the NFR Framework notation is still being determined.

Most RE experts agree that the catalogue is simple to understand, but they attribute this ease of comprehension to their familiarity with the NFR Framework's notation rules. On the other hand, HRI experts approve of the information in the catalogue but disapprove the use of NFR Framework notation. Our initial hypothesis was that using the NFR Framework would enable a more structured, concise, and easily understood representation of information. While this holds for the Requirements Engineering community, it does not have the same effect in the Human-Robot Interaction community. Further investigation into alternative approaches or tools for organizing information in catalogues may be relevant.

Due to the limited number of participants in the evaluation, the conducted assessment served as a preliminary evaluation. Nonetheless, this initial evaluation yielded valuable insights into the challenges that need to be addressed in future validation processes. These insights will be instrumental in refining and improving the validation methodology for subsequent evaluations.

G. Assessment by Requirements Engineers

We invited the participants of this evaluation in two stages: first via e-mail to the participants who are members of a requirements engineering community. The second through activity in Google Classroom with students from graduate requirements engineering course. As a result, we were able to engage 20 volunteers RE participants.

We achieved this remotely on videoconference, supported by a questionnaire in Google Forms. The profiles of the participants in this research consisted of Requirements Engineers, most of them with a good knowledge of requirements elicitation and the NFR Framework. However, only 40% with a good understanding of Human-Robot Interaction and relatively low knowledge of the other areas (Trust, Social Robotics, and Socially Assistive Robotics).

The results obtained by the questionnaire application allowed us to conclude that the requirements presented in the Catalogue reflect, with reasonable confidence, some of the qualities of Trust. The participants positively validated the pertinence of these requirements. It was also possible to check that the defined correlations satisfactorily reflected the impacts of one NFR on another.

According to the participants, the catalogue presented a suitable range of Non-Functional Trust Requirements in Socially Assistive Robotics. We included most of the Trust factors suggested or highlighted as important by the participants in the catalogue. Some participants made observations about the requirements presented, like including new ones and minor modifications.

As for correlations, most participants agreed that they were adequate. Participants also suggested some modifications to improve the catalogue. We balanced these considerations, incorporating some of the suggestions for the current version of the catalogue while saving other aspects for future revisions. More details are available at [25].

VII. RELATED WORKS

In the literature, we found no catalogues specifically targeting Trust as a Non-Functional Requirement for Socially Assistive Robots. However, some existing works investigate Trust as an NFR in different domains [35, 36]. Nonetheless, these works are not directly related to the Human-Robot Interaction field. We based our research on other works that developed NFR catalogues in various domains [37, 38, 39]. Our work represents another application case of the original SIG methodology.

Our taxonomy is influenced by key works such as [20, 21, 22, 23]. Tables IV and V provide an overview of the nine relevant works that influence our study.

TABLE IV. Comparison of related works on observed aspects

Works	*Has Taxonomy?*	*It's about Trust?*	*It's about HRI?*	*It's about SARs?*	*Deals with Safety/Privacy in Trust?*
Hancock, Peter A., et al. (2011) [20]	Yes	Yes	Yes	No	No
Schaefer, Kristin E. (2013) [21]	Yes	Yes	Yes	No	No
Schaefer, Kristin E., et al. (2016) [22]	Yes	Yes	No	No	No
Langer, Allison, et al. (2019) [23]	No	Yes	Yes	Yes	Yes
NFR4TRUST [25]	Yes	Yes	Yes	Yes	Yes

TABLE V. Comparison of related works on NFR catalogues

Works	*Presents catalogues with the NFR Framework*	*It's about Trust?*	*It's about Trust in HRI?*	*It's about NFRs for SARs?*
Cysneiros, L. M., do Prado Leite, J. C. S. (2020) [35]	Yes	Yes	No	No
Kwan, D., Cysneiros, L. M., do Prado Leite, J. C. S. (2021) [36]	Yes	Yes	No	No
Silva, R. A. D (2019) [37]	Yes	No	No	No
Quintanilla Portugal, Roxana Lisette. (2020) [38]	Yes	No	No	No
Sadi, Mahsa Hasani. (2020) [39]	Yes	No	No	No
NFR4TRUST [25]	Yes	Yes	Yes	Yes

VIII. THREATS TO VALIDITY

A study's validity lies in assessing how truthful and uninfluenced by researchers' biases the research results are. It is crucial to define potential threats to this trustworthiness. For our catalogue, we considered the following threats [35]:

Threats to Constructor Validity: These threats can affect the accuracy and reliability of the measurement of theoretical concepts in a study.

1. **Hypothesis guessing:** This threat deals with the effect caused by participants in a study becoming aware of the researcher's hypotheses or expectations, influencing their responses or behaviour in a way that aligns with those expectations. We kept participants unaware of the assumptions or conditions to mitigate the threat.
2. **Social desirability bias:** This threat deals with how participants may provide responses that they believe are socially desirable, leading to distorted or biased data. To mitigate this threat, we guarantee participants the confidentiality of their responses. We also observed the participants' behaviour through recorded materials and used neutral and non-inducing language, avoiding excessively positive or negative phrases during the interviews and questionnaire application.
3. **Participant characteristics:** This threat deals with how differences among participants, such as their demographic background, prior experiences, or personal biases, may affect their responses or behaviour in ways that impact the validity of the measures. To mitigate this threat, we provide explicit instructions and guidelines to participants to minimize the influence

of personal biases or prior experiences on their responses.

Threats to Internal Validity: These threats can affect the relationship between treatment and outcome.

1. **Maturation:** This threat deals with how subjects behave over time. To mitigate this threat, we limited the interview time and the content needed to understand the topic.
2. **Instrumentation:** This threat deals with the effect caused by the artifacts used in the research. Given the multidisciplinary nature of our research, we adapted the instruments used according to the participants' general profiles.

Threats to External Validity: These threats can affect the generalization of the experiment result to other environments.

1. **Selection interaction and treatment:** This threat deals with the effect caused by having subjects not representative of the population. In part of the evaluation of this research, we used students as subjects. However, we reduced this threat by using graduates who study or have experience in requirements engineering. Another threat to the validity of this type is the selection of specialists for evaluation who may not have the necessary knowledge of some multidisciplinary aspects of the research. To mitigate this threat, we make a brief presentation of the main concepts according to the knowledge profile of each specialist.

Threats to Conclusion Validity: These threats can affect the ability to draw correct conclusions about the study results.

1. **Low Statistical Power:** This threat indicates that low statistical power induces a high risk of drawing an erroneous conclusion. Due to the low number of experts with good knowledge in all the main aspects (RE, HRI, TRUST, SARs), there may not be enough observations in the study, which can lead to inappropriate conclusions. We accept the threat, justified by the number of experts participating, given that there are problems with the incompatibility of schedules and time. We tried to mitigate the low number of participants, looking for specialists with more experience in your specific domain. Related to expert knowledge, we tried to mitigate it by knowing in advance the aspects that the experts lacked while presenting these aspects with more emphasis.
2. **Fishing:** This threat deals with the influence of the researchers in the results by searching for a specific outcome. We tried to mitigate this threat by using the Likert scale on some questions and modelling the most addition, we conducted these meetings with two different interviewers (together or not) who followed the same interview protocol.
3. **Heterogeneity of Subjects:** This threat deals with the effects of a significantly heterogeneous group (or participants). We tried to mitigate this threat by assessing each participant's knowledge of the discussed topics and presenting with more emphasis the fundamentals they lacked.

IX. Conclusions and Future Work

This paper focused on identifying, examining, and documenting the Non-Functional Requirements of Trust specific to Anthropomorphic-type Socially Assistive Robots. The research questions were successfully addressed, with the identification and documentation of 125 Trust NFRs that can be of importance for the design and selection of SARs as well as the development of their applications. Several (14) primary trust requirements were identified (see Figure 1). These NFRs play a critical role in establishing trust and ensuring the effective functioning of SARs.

Regarding the use of the NFR Framework notation for the representation of the Trust Catalogue, preliminary feedback from experts in Human-Robot Interaction and Requirements Engineering was mixed. While the RE community found it beneficial, HRI experts expressed concerns about its applicability. This highlights the need to explore alternative structuring approaches, better suited to HRI contexts, ensuring an effective and accessible representation of trust-related catalogues.

Note that the effort required to utilize the catalogue and its connections to real requirements is not explicitly clear. This aspect should be addressed in future research to provide guidance on how engineers can effectively leverage the catalogue and apply its definitions and templates to real-world scenarios.

Furthermore, future directions include expanding the scope of expert selection to include individuals with expertise in both Human-Robot Interaction and Requirements Engineering, ensuring a more comprehensive understanding of the subject matter. Conducting a systematic literature review to enhance the Trust taxonomy, validating it through experiments, expanding the catalogue with additional trust factors, simplifying it for better lay user accessibility, applying the catalogue to diverse SAR applications beyond physical therapy, and establishing a repository for managing non-functional trust requirements. These steps will contribute to advancing research and development in Anthropomorphic-type Socially Assistive Robots, facilitating a deeper understanding and implementation of trust-related requirements in HRI.

Acknowledgment

The authors would like to express their gratitude for the financial support provided by Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (**FACEPE**), including Proj. **IBPG-0370-1.03/22** and Proj. **PRONEX-APQ-0880-1.03/14**, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (**CAPES**), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (**CNPq**). Their support has been instrumental in the successful completion of this research.

References

[1] V. P. Binotte, "Desenvolvimento de um robô socialmente assistivo com controle baseado em comportamento de seleção de ação para interação com crianças com TEA", Master's thesis, Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Espírito Santo, Brazil, 2018.

[2] D. Feil-Seifer and M. J. Matarić, "Socially Assistive Human-Robot Intervention for Children with Autism Spectrum Disorders," in 11th Int. Symp. on Experimental Robotics (ISER), vol. 54, pp. 201-210, 2008.

[3] B. Scassellati, H. Admoni, and M. Matarić, "Robots for use in autism research", Annual review of biomedical engineering, vol. 14, pp. 275–294, 2012.

[4] H. Yan, M. H. Ang, and A. N. Poo, "A survey on perception methods for human-robot interaction in social robots" , International Journal of Social Robotics, vol. 6, no. 1, pp. 85–119, 2014.

[5] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance", Human factors, vol. 46, no. 1, pp. 50–80, 2004.

[6] N. Luhmann, " Trust and power Chichester", UK: Wiley, 1979.

[7] P. Kollock, "The emergence of exchange structures: An experimental study of uncertainty, commitment, and trust," American Journal of sociology, vol. 100, no. 2, pp. 313-345, 1994.

[8] N. Luhmann, "Familiarity, confidence, trust: Problems and alternatives," Trust: Making and breaking cooperative relations, vol. 6, no. 1, pp. 94-107, 2000.

[9] M. Salem and K. Dautenhahn, "Evaluating Trust and Safety in HRI: Practical Issues and Ethical Challenges," in Emerging Policy and Ethics of Human-Robot Interaction: A Workshop at 10th ACM/IEEE Int Conf on Human-Robot Interaction (HRI 2015), ACM Press, Portland, United States, March 2, 2015.

[10] A. Rossi, K. Dautenhahn, K. L. Koay, and J. Saunders, "Investigating human perceptions of trust in robots for safe HRI in home environments," in Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, 2017, pp. 375-376.

[11] A. Freedy, E. DeVisser, G. Weltman, and N. Coeyman, "Measurement of trust in human-robot collaboration," in 2007 international symposium on collaborative technologies and systems, 2007, pp. 106-114.

[12] L. Chung, B. A. Nixon, E. Yu, and J. Mylopoulos, "Non-functional requirements in software engineering," vol. 5. Springer Science & Business Media, 2000.

[13] M. A. Goodrich and A. C. Schultz, "Human-robot interaction: A survey," Foundations and Trends in Human-Computer Interaction, vol. 1, no. 3, pp. 203-275, 2008.

[14] G. Charalambous, S. Fletcher, and P. Webb, "The development of a scale to evaluate trust in industrial human-robot collaboration," International Journal of Social Robotics, vol. 8, no. 2, pp. 193-209, 2016.

[15] P. A. Lasota, T. Fong, and J. A. Shah, "A Survey of Methods for Safe Human-Robot Interaction," Foundations and Trends in Robotics, vol. 5, no. 3, pp. 261-349, 2017.

[16] K. Blomqvist, "The many faces of trust," Scandinavian Journal of Management, vol. 13, no. 3, pp. 271-286, 1997.

[17] D. Mairiza, D. Zowghi, and N. Nurmuliani, "An investigation into the notion of non-functional requirements," Proceedings of the 2010 ACM Symposium on Applied Computing - SAC '10, 2010, DOI: https://doi.org/10.1145/1774088.1774153.

[18] K. Khatter and A. Kalia, "Impact of Non-functional Requirements on Requirements Evolution," IEEE Xplore, Dec. 01, 2013.

[19] L. R. Costa, J. Castro, J. Kelner, M. Lencastre, and O. Pastor, "On the Use of Social Robots for rehabilitation: The case of NAO Physio," in Proceedings of the 6th International Conference on Information Technology & Systems (ICITS'23), vol. 2, Lecture Notes in Networks and Systems (LNNS), 2023, ISBN 978-3-031-33260-9.

[20] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," Human factors, vol. 53, no. 5, pp. 517-527, 2011.

[21] K. Schaefer, "The Perception And Measurement Of Human-robot Trust," University of Central Florida, USA, 2013.

[22] K. E. Schaefer, J. Y. C. Chen, J. L. Szalma, and P. A. Hancock, "A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems," Human Factors, vol. 58, no. 3, pp. 377-400, 2016.

[23] A. Langer, R. Feingold-Polak, O. Mueller, P. Kellmeyer, and S. Levy-Tzedek, "Trust in socially assistive robots: Considerations for use in rehabilitation," Neuroscience and Biobehavioral Reviews, vol. 104, no. July, pp. 231-239, 2019.

[24] D. Feil-Seifer and M. J. Matarić, "Defining socially assistive robotics," Proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics, vol. 2005, pp. 465-468, 2005.

[25] L. R. Costa, "NFR4TRUST: Catálogo de Requisitos Não- Funcionais de Confiança para Robôs Socialmente Assistivos" (in English, NFR4TRUST: Non-Functional Requirements Catalogue for Socially Assistive Robots). Master's thesis. Universidade Federal de Pernambuco. Brazil, 2022.

[26] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," Robotics and autonomous systems, vol. 42, no. 3-4, pp. 143-166, 2003.

[27] B. R. Duffy, "Anthropomorphism and the social robot," Robotics and autonomous systems, vol. 42, no. 3-4, pp. 177-190, 2003.

[28] C. Bartneck, T. Kanda, O. Mubin, and A. Al Mahmud, "Does the design of a robot influence its animacy and perceived intelligence?," International Journal of Social Robotics, vol. 1, no. 2, pp. 195-204, 2009.

[29] I. Papadopoulos, R. Lazzarino, S. Miah, T. Weaver, B. Thomas, and C. Koulouglioti, "A systematic review of the literature regarding socially assistive robots in pre-tertiary education," Computers & Education, vol. 155, p. 103924, 2020.

[30] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behavior to tasks to improve human-robot cooperation," in The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003, pp. 55-60, IEEE.

[31] V. K. Sims et al., "Anthropomorphism of Robotic Forms: A Response to Affordances?," vol. 49, no. 3, pp. 602-605, Sep. 2005, doi: https://doi.org/10.1177/154193120504900383.

[32] D. Feil-Seifer and M. J. Matarić, "Socially Assistive Human-Robot Intervention for Children with Autism Spectrum Disorders," in 11th Int. Symp. on Experimental Robotics (ISER), vol. 54, pp. 201-210, 2008.

[33] C. McNamara. "General Guidelines for Conducting Research Interviews," Free Management Library, "Wording of Questions" section, para. 1. sept. 27, 2022. https://managementhelp.org/businessresearch/interviews.htm.

[34] D. Dunning, "The Dunning--Kruger effect: On being ignorant of one's own ignorance," in Advances in experimental social psychology, vol. 44, Elsevier, 2011, pp. 247-296.

[35] L. M. Cysneiros and J. C. S. do Prado Leite, "Non-functional requirements orienting the development of socially responsible software," in Enterprise, Business-Process and Information Systems Modeling, Springer, 2020, pp. 335-342.

[36] D. Kwan, L. M. Cysneiros, and J. C. S. do Prado Leite, "Towards Achieving Trust Through Transparency and Ethics," in 2021 IEEE 29th International Requirements Engineering Conference (RE), 2021, pp. 82-93.

[37] R. A. Silva, "NFR4ES: Um catálogo de requisitos nao-funcionais para sistemas embarcados." Master's thesis. Universidade Federal de Pernambuco, 2019.

[38] R. L. Quintanilla Portugal, "Speeding-Up Non-Functional Requirements Elicitation," Ph.D. thesis, Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2020.

[39] M. H. Sadi, "Assisting with API Design through Reusing Design Knowledge," Ph.D. thesis, University of Toronto, Canada, 2020.

[40] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, Experimentation in software engineering. Springer Science & Business Média, 2012.