"Socially Assistive Robot Privacy Model": A Multi-model Approach to Evaluating Socially Assistive Robot Privacy Concerns

Sawyer Collins[1(✉)], Čedomir Stanojević[2], Casey Bennett[3], Zachary Henkel[4], Kenna Baugus Henkel[4], Nikki M. Abbott[2], Cindy L. Bethel[4], and Selma Šabanović[1]

[1] Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, Bloomington, IN, USA
sercolli@iu.edu
[2] College of Behavioral, Social, and Health Sciences, Clemson University, Clemson, SC, USA
[3] College of Computing and Digital Media, DePaul University, Chicago, IL, USA
[4] Department of Computer Science and Engineering, Mississippi State University, Starkville, MS, USA

Abstract. As socially assistive robots (SARs) enter more diverse care settings, including users' homes, it is critical to identify the shifting privacy risks and dimensions of privacy this technology and its data collection capabilities may affect. We propose a new model of privacy to address the complex nature of SARs as a multimodal technology within the healthcare space. To construct this new model, we combine a previous three-dimensional model from the healthcare literature and synthesize it with a seven-dimension technology-related model. We then use this new combined model to analyze self-report data from several workshops with prospective users of the dog-like robot Therabot to map out the dimensions of privacy identified as future concerns for clinicians and those living with depression. Finally, we suggest this model can be used in future studies to support the in-depth exploration of privacy implications of SARs within healthcare through discussions about privacy among clinicians, those receiving care, and robot designers.

Keywords: Socially Assistive Robots · Healthcare Privacy · Social Robot Privacy

1 Introduction

Privacy and healthcare are critically linked - from patients' perceptions of who has admissible access to their health information to when that information can or should be accessed. Privacy of patient information can even change their stance toward the healthcare services that they are willing to receive [15]. Given that, healthcare technologies need to adhere to strict standards to protect the

O. Palinko et al. (Eds.): ICSR + AI 2024, LNAI 15561, pp. 280–289, 2025.
https://doi.org/10.1007/978-981-96-3522-1_25

individual's right to privacy. The privacy implications of socially assistive robots (SARs), which exist to provide social support to the individual user, particularly in areas such as mental health [7], are still underexplored. The use of SARs in mental healthcare has expanded from laboratory testing to implementing robots in more personal settings such as nursing homes [2] and in user's own homes [18]. SARs can provide unique insights into these environments, including monitoring the status of the individual in their home through onboard sensors. It is therefore both timely and necessary to evaluate the effects of these robots on individuals' perceptions of privacy. This includes questions of understanding who has access to the data that is collected by the robot, when they have access to that data, and what is done with that information (such as how it is stored, processed, and shared).

SARs intersect both the technology and healthcare domains, so identifying the different aspects of an individual's right to privacy that their use may put at risk requires reflection on privacy in both domains. In this paper, we propose a comprehensive privacy risk model, which we refer to as the Socially Assistive Robot Privacy Model (SARPM), which combines the insights of several studies to identify the effects of trust on patient care and to further break down the different privacy aspects for a given technology, and provides context within the healthcare domain for understanding the interactions with the data collected type. Using this model in future studies on SARs in healthcare will further interdisciplinary discussions with users, researchers, and the larger care team surrounding privacy concerns with these technologies.

2 Privacy Background and SARPM

2.1 Privacy and Healthcare

Privacy in healthcare has been described as "the ability of an individual or group to stop information about themselves from becoming known to people other than those they choose to give the information to" [11]. Within the healthcare sector, three main dimensions of privacy have been established by Serenko et al.: informational, physical, and psychological [19]. Informational privacy includes data that the patient or client may disperse to their physician or therapist about their health needs, and includes big data collection across multiple platforms and providers (i.e. patient health data, progress notes) [20]. Physical privacy includes how available one is to others, and in this context can include items that collect patient's data more frequently throughout the day (i.e. use of wearable devices continually touching user) [3]. Psychological privacy includes information regarding an individual's thoughts and feelings, and how those thoughts are respected by their care team [19]. Each of these dimensions are important to provide the individual both with adequate care as well as support their trust in adopting technologies in healthcare.

2.2 Privacy and Social Robots

Beyond being known for providing social and emotional support, social robots have been used to help monitor activity within the users' homes [14]. Regular use of the robot in this way can produce situations in which users may feel their privacy is violated, due to information collected by the robot that may become available to others, including the person's caregivers, friends and family members, and third parties. This specific area furthermore provides a unique and potentially worrying set of concerns that could violate one's privacy [5]. Social robots are not immune to potential hacking by unwanted entities and thus security of the robotic sensor suites should be a priority, lest the user fall victim to spying. Perceptions of the cybersecurity and privacy risks influence user comforts and trust in adopting such technologies [12].

This unique aspects of social technologies is recognized within the robotics community, such as by Lutz et al. who identified four dimensions of privacy within the context of social robots recognizing the differences between robots who work as social actors and technologies such as smart phones [13]. Lutz et al. identified physical, social, psychological, and informational, recognizing that social robots provide the opportunity for further attachment of the user stressing the importance of privacy beyond just informational [13]. Further, privacy has been discussed within the context of telepresence robots, including the testing of both humanoid and mobile robots [16]. This testing provides helpful insight into the explainability of the robot to display when the robot is collecting data, versus when it is not, display physical changes to indicate the stopping of specific sensors based upon the individuals privacy concerns [16].

One model of privacy that was originally developed to provide deeper context for the different privacy dimensions of technology has been explored in previous research within the context of social robots as a whole [10]. That model included seven dimensions of privacy that can be affected by technology: privacy of person; action and behavior; communication; data and images; thoughts and feelings; location and space; and association [8]. Privacy of person is the ability to control ones physical data, privacy of action and behavior is the ability to withhold information regarding the physical behaviors of the individual, privacy of communication includes things such as phone calls or emails, privacy of data and images implies being able to withhold items such as photos, privacy of thoughts and feelings is the ability to withhold opinion, privacy of location and space is the ability to withhold information regarding tracking information, and finally privacy of association is the ability to withhold information regarding personal connections with others [8]. This model has been previously used with other technologies, including whole body scanners [8], evaluating drones [21], and critically exploring smart cities [6]. Through the seven dimensions, researchers are able to best identify the areas of increased concern for their technologies based upon use case, but it does not provide guidance within the specific context of healthcare.

3 SARPM: Socially Assistive Robot Privacy Model

The Socially Assistive Robot Privacy Model works through identifying both a dimension from the original healthcare model, as well as a corresponding technology dimension. Through the use of two previously described models, we have constructed a more explicit model which details privacy considerations a health provider should take into account when utilizing the social robot as a therapeutic tool. If a health practitioners was to follow this model, they would be cognizant of their patient's privacy needs (i.e. psychological privacy domain) and structure their communication, with a patient by placing special attention to patient's thoughts and feelings, and association. The dimensions identified from the technological privacy domains breakdown how the robot may intrude on an individuals privacy, while the relation to psychological privacy from the healthcare domain helps provide a clearer path to handling that information by the care team and those implementing the robot.

This model (Fig. 1) may be used in the future to further explain the different aspects of privacy that users and clinicians are most concerned about and the effects of the different aspects of SARs on an individual's right to privacy. Furthermore, the SARPM can be used as a guiding tool for SAR designers to understand what privacy aspects of novel technology need to be taken into consideration, and provide suggestions to the link between the healthcare dimensions of privacy and technology dimensions of privacy.

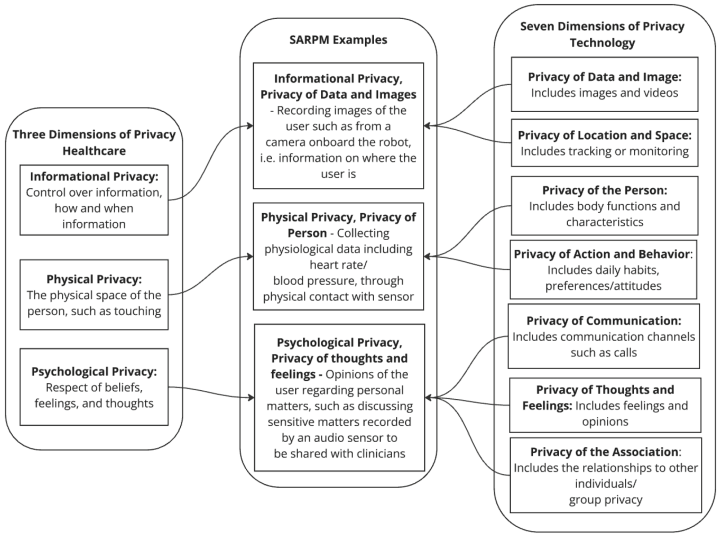

Fig. 1. Healthcare privacy dimensions, technology privacy dimensions, and SARPM examples

It can be used as the informational tool for users of SARs on how their privacy is being handled, since it is straightforward enough for non-industry and people outside of academia to comprehend. Aspects of explainability are notoriously difficult when utilizing new technology, thus allowing a detailed breakdown of the effects within the context of technology itself as well as within the dimensions of healthcare may allow a more grounded discussion on future privacy implications between designers and those they are designing for.

4 Developing Privacy Dimensions Through User Studies with Therabot

Socially assistive robots are unique in that they exist as social actors in both the domains of technology and healthcare. Finding a privacy model that fully encompasses all aspects of that kind of social technology can be difficult, as social interaction with humans is by definition more personal in nature. To attempt to mitigate this and fully explore the privacy dimensions within the context of SARs as a healthcare device in mental healthcare, we propose looking through the combined lens of the two models mentioned above: the three-dimension model used within healthcare [19] and the seven-dimension model used for technology [8,10]. The resulting SARPM: Socially Assistive Robot Privacy Model was formed after an initial workshop was conducted regarding privacy concerns and the robot Therabot, further suggesting the need for identifying dimensions for discussing specific privacy concerns amongst various stakeholders. Focusing on the robot Therabot (Fig. 2), which is a socially therapeutic assistive robot designed to support mental health, we developed and applied the proposed model. This robot is shaped like a stuffed-animal dog, and it is intended to provide comfort and companionship through interactive behaviors.

Fig. 2. Therabot Robot

Study one which inspired the creation of SARPM was an exploratory design evaluation study conducted in the spring of 2019 at the University of Technology, Sydney, to investigate the use of Therabot as reported in [9]. Following

an interview regarding anxiety and stress, all participants interacted with the robot during an interactive design portion of the session where they provided feedback on Therabot's design and behaviors followed by their own preferences and privacy considerations. The results of this study suggested a need for the consideration of privacy within the context of SARs use and design, particularly within the context of specific actionable dimensions, leading to the development of the SARPM.

In study two we conducted two independently run online workshops involving individuals living with depression focused on the design of a personalized Therabot robot [4]. In the first workshop, five individuals with a confirmed depression diagnosis participated in a series of five workshop sessions. In the second workshop, ten individuals who self-reported having a depression diagnosis participated in a single one-hour session. Participants were asked to choose a physical design for the robot, discuss where and how it would be used, and what sensors the robot should have, among other various aspects of the robots design. All participants were also asked to consider privacy related topics when discussing data collection with the Therabot robot, such as using the data collected in therapy or with their care team, and from this information the model was used to identify their perceptions regarding privacy concerns.

In the third and final study, four clinicians who work with those living with depression discussed designs for Therabot during multiple online sessions. Clinicians were first introduced to SARs through videos of Buddy [1] and Paro [17], and gave initial feedback regarding how these robots in general may be used with their populations. In later workshop sessions they were asked to review the designs of Therabot specifically as created by those living with depression. They gave their feedback on the physical design, sensors, behaviors, and uses of the Therabot. Clinicians also spoke about how they might use the robot in their own practice, and what aspects they believed would be most beneficial [4]. From these discussions the SARPM was applied to identify which dimensions of privacy those living with depression and clinicians were most concerned with.

4.1 Study 1: Initial Privacy Perceptions

In our initial study ($n = 27$) [9], participants emphasized the importance of informed consent and transparency, with some addressing these areas directly prior to being explicitly asked. They discussed personal information and expressed preferences to be aware of exactly what data the robot collected as well as when and with whom it would be shared. Visibility and control over data were also key concerns. Participants wanted clear and easy ways to see who had access to their data and to manage this access themselves, helping them feel more secure and in control.

Participants generally expressed comfort in sharing robot usage logs with their therapists ($n = 27$). When asked open-ended questions to identify the most useful features for therapists in a robot companion app, participants suggested usage information ($n = 13$), robot settings ($n = 5$), biometric health data ($n = 3$), and alerts of user distress ($n = 2$). In addressing user privacy concerns,

participants emphasized the importance of understanding data usage, including what data is collected, how it is used, and who has access to it ($n = 10$). They also stressed the need for consent from all present parties ($n = 9$), and the importance of the robot not persistently recording audio ($n = 4$). In this study, participants were made aware by the researcher that Therabot does not contain a camera for privacy purposes; however, 2 participants emphasized that video recording would be a privacy concern if it were added.

Overall, these privacy findings highlight the need for personal robots to be designed with robust privacy protections and user-friendly controls to ensure users are in control of their personal information. Further, this initial study suggested the need for identification of the specific dimensions regarding privacy concerns related to the robot. To accommodate this as studies regarding specific design recommendations for Therabot progressed, the SARPM was developed, and later applied to continuing participatory design responses.

4.2 Study 2: Privacy Dimensions for Users

During the workshops with those living with depression [4], many participants were interested in the robot's ability to sense touch (n = 12), with many indicating that they would touch the robot when they needed support or care. By desiring interaction with the robot in this way, participants are both indicating the need for specific identification of the type of privacy that could be affected, such as physical privacy of thoughts and feelings (such as being able to control when the robot is touching them), as well as through informational privacy of thoughts and feelings through the reporting of this data to their care team, particularly in the case of emergencies. Participants also requested the robot to collect auditory information. They requested that Therabot be able to recognize keywords (n = 5) such as its name but not continuously record conversations. This focus on the robot's inability to record continuous conversations points to a consideration of psychological privacy in communication. Physiological sensors that could monitor aspects such as heart rate and blood pressure were also of interest to them. This would follow the physical privacy of the person, as many of these sensors would have to be worn.

Overall participants were interested in the robot having some ability to visually recognize that they were present (n = 9). However, most participants were not interested in Therabot being equipped with a typical camera (n = 8) emphasizing apprehension that is aligned with informational privacy of data and images dimension of SARPM. According to the participants, security of images and a risk of a visual data leak is their biggest concern that could invade their privacy. This presents a new privacy concern, particularly that of informational privacy of location and space - as a visual sensor of this sort may allow the robot to know where the individual is within the home.

Through the context of the discussion of these sensors with participants and applying the SARPM, the primary dimensions of concern could potentially be are physical privacy of thoughts and feelings, physiological privacy of communication, physical privacy of person, and informational privacy of data and images.

4.3 Study 3: US Clinicians

Overall, clinicians who worked with those living with depression were most interested in the robot's ability to collect physiological data about their client, such as heart rate and blood pressure (n = 3), as indicated by those living with depression [4]. The informational privacy of person aspect appears both within the context of those living with depression and clinicians. While those living with depression were more interested in the robot being able to collect auditory information for aspects such as knowing their name, clinicians were more interested in tell-tale markers, including aspects of emotional changes such as sighing or crying (n = 3). Due to the nature of the type of information being requested, rather than the psychological privacy of communication, this becomes informational privacy of thoughts and feelings through the expression of emotion. Again, in regard to privacy concerns, a camera was seen as something that may prevent the clinician's clients from utilizing the robot (n = 2). Their concern was the informational privacy of data and images, voicing that their clients may feel that the images could be leaked or monitored by unwanted entities.

Clinicians primarily focused on the informational privacy dimension due to the collection of others' data that would be shared with them, though two clinicians mentioned concern about robots collecting their information as well. Sensors such as physiological sensors, auditory sensors, and visual sensors were of interest or concern to the clinicians. Through further discussion the dimensions of privacy identified are: informational privacy of person, informational privacy of thoughts and feelings, and informational privacy of data and images.

5 Discussion

Socially assistive robots are complex in that they provide multi-modal interactions as well as the potential for data collection and insight into the users condition. While this complexity can deepen interactions, and provide a reflection into the users' daily life with their care team, it also poses a risk to the individual's privacy. Due to the complex nature of SARs, it is critical that they are assessed not just as technology, but specifically include the dimensions and considerations relevant for healthcare privacy. Through our workshops with individuals assessing a SAR prototype Therabot, we identified the need for a SARPM that would allow all stakeholders involved in the design and use of SAR to discuss the complex dimensions of privacy. This model combines two pre-existing models, one from healthcare [19] and one from technology [8], to create a more comprehensive approach to privacy for healthcare technology.

Through the use of this SARPM, we mapped specific dimensions of privacy to requests by those living with depression and clinicians about what data they would like Therabot to collect regarding themselves and their environment. These dimensions from those living with depression (study two): physical privacy of thoughts and feelings, physiological privacy of communication, physical privacy of person, and informational privacy of data and images. In comparison, clinicians primarily focus on informational privacy of person, informational

privacy of thoughts and feelings, and informational privacy of data and images. This discrepancy between the end user and cursory stakeholder further shows the importance of the SARPM in the discussion between all stakeholders in developing SARs. Using the model to map the privacy concerns surrounding the robot and specific aspects will allow for a clear, detailed explanation of the privacy dimensions affected at each level and encourage continuous discussion as new dimensions emerge. While clinicians (study three) focused entirely on the informational dimensions, this suggests a potential lack of scrutiny for their clients' psychological or physical implications as the robot is designed, whereas those living with depression indicated a much wider range of considerations for the robot.

By using the SARPM to break down the specific dimensions within the different aspects of SARs, transparency regarding data collection, how the data may be used, and to whom the data is being stored can be encouraged. By providing connections to both previous technological privacy domains and healthcare domains, SARPM may aid in the handling and processing of data collected by the robot - particularly important as the robot works to encourage social interaction. Further, this model provides a template for those involved in the development and use of SARs to identify the impact of the data collection directly on healthcare information and further narrow the specific risks influenced by the design of the robot. Encouraging this discussion amongst all stakeholders early in the development of SARs can help minimize the risks to the user's privacy by ensuring that all involved understand what information is being collected and how it is being used; through this combined model, we aim to support these discussions.

6 Conclusion

Combining both a three dimension model of privacy in healthcare, and a seven dimension model of privacy for technology, we propose the Socially Assistive Robot Model as a tool to expand the explainability and analysis of dimensions of privacy for SARs. Our broader aim is to help other researchers by providing a more comprehensive framework for privacy specifically in the healthcare technology space, including robots, wearables, and other IOT devices.

References

1. Buddy: Your family's companion robot (2024). https://www.bluefrogrobotics.com/robot
2. Chang, W.L., Šabanović, S.: Studying socially assistive robots in their organizational context: Studies with paro in a nursing home. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts 10(1), pp. 227–228 (2015)
3. Cilliers, L.: Wearable devices in healthcare: privacy and information security issues. Health Inf. Manage. J. **49**(2), 150–156 (2020)

4. Collins, S., et al.: An emotional support animal, without the animal: design guidelines for a social robot to address symptoms of depression. In: Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction, pp.147–153 (2024)
5. Denning, T., Matuszek, C., Koscher, K., Smith, J.R., Kohno, T.: A spotlight on security and privacy risks with future household robots: attacks and lessons. Proceedings of the 11th International Conference on Ubiquitous Computing, pp. 105–115, January 2009
6. Eckhoff, D., Wagner, I.: Privacy in the smart city—applications, technologies, challenges, and solutions. IEEE Communications Surveys Tutorials, pp. 489–516, March 2017
7. Feil-Seifer, D., Mataric, M.J.: Defining socially assistive robotics. In: 9th International Conference on Rehabilitation Robotics, pp. 465–468, March 2005
8. Finn, R.L., Wright, D., Friedewald., M.: Seven types of privacy. European data protection: coming of age, pp. 3–32, March 2013
9. Henkel, K., Henkel, Z., Aldridge, A., Bethel, C.L.: The evolving design of a socially therapeutic robotic dog*. In: 2023 World Symposium on Digital Intelligence for Systems and Machines (DISA), pp. 90–97 (2023). https://doi.org/10.1109/DISA59116.2023.10308940
10. Heuer, T., Schiering, I., Gerndt., R.: Privacy-centered design for social robots. Interaction Stud. **20**(3), 509–529 (2019)
11. Laric, M.V., Pitta, D.A., Katsanis., L.P.: Consumer concerns for healthcare information privacy: A comparison of us and Canadian perspectives. Res. Healthcare Financ. Manage. **12**(1), 03 (2009)
12. Levinson, L., McKinney, J., Nippert-Eng, C., Gomez, R., Šabanović., S.: Our business, not the robot's: family conversations about privacy with social robots in the home. Front. Robot. AI **11**, 1331347 (2024)
13. Lutz, Christoph, M.S., Hoffmann., C.P.: The privacy implications of social robots: scoping review and expert interviews. Mob. Media Commun. **7**(3) (Mar 2019)
14. Napoli, D., Claudia, G.E., Rossi., S.: Personalized home-care support for the elderly: a field experience with a social robot at home. User Modeling and User-Adapted Interaction **33**(2), 405–440 (2023)
15. Nayeri, N.D., Aghajani, M.: Patients' privacy and satisfaction in the emergency department: a descriptive analytical study. Nurs. Ethics **17**(2), 167–177 (2010)
16. Nieto Agraz, Celia, P.H.M.E., Hein, A.: Is the robot spying on me? a study on perceived privacy in telepresence scenarios in a care setting with mobile and humanoid robots. Int. J. Soc. Robot., 1–15, March 2024
17. Paro (2014). http://www.parorobots.com/
18. Randall, N., Bennett, C.C., Šabanović, S., Nagata, S., Eldridge, L., Collins, S., Piatt, J.A.: More than just friends: in-home use and design recommendations for sensing socially assistive robots (sars) by older adults with depression. Paladyn, J. Behav. Robot. **10**(1), 237–255 (2019)
19. Serenko, N., Fan, L.: Patients' perceptions of privacy and their outcomes in healthcare. Int. J. Behav. Healthcare Res. **4**(2), 101–122 (2013)
20. Sippe, R.: The relation between big data and informational privacy in the context of the healthcare. Masters Thesis, January 2015
21. Yaacoub, J.P., Noura, H., Salman, O., Chehab, A.: Security analysis of drones systems: attacks, limitations, and recommendations. Internet Things **11**, 100218 (2020)