
Final accepted version (with author's formatting)

This version is available at: http://eprints.mdx.ac.uk/24838/

Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this work are retained by the author and/or other copyright owners unless otherwise stated. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge.

Works, including theses and research projects, may not be reproduced in any format or medium, or extensive quotations taken from them, or their content changed in any way, without first obtaining permission in writing from the copyright holder(s). They may not be sold or exploited commercially in any format or medium without the prior written permission of the copyright holder(s).

Full bibliographic details must be given when referring to, or quoting from full items including the author’s name, the title of the work, publication details where relevant (place, publisher, date), pagination, and for theses or dissertations the awarding institution, the degree type awarded, and the date of the award.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:

eprints@mdx.ac.uk

The item will be removed from the repository while any claim is being investigated.

See also repository copyright: re-use policy: http://eprints.mdx.ac.uk/policies.html#copy
Embedding Ethics in the Design of Culturally Competent Socially Assistive Robots

L. Battistuzzi, C. Papadopoulos, I. Papadopoulos, C. Koulouglioti, A. Sgorbissa.

Abstract— Research focusing on the development of socially assistive robots (SARs) for the care of older adults has grown in recent years, prompting a great deal of ethical analysis and reflection on the future of SARs in caring roles. Much of this ethical thinking, however, has taken place far from the settings where technological innovation is practiced.

Different frameworks have been proposed to bridge this gap and enable researchers to handle the ethical dimension of technology from within the design and development process, including Value Sensitive Design (VSD). VSD has been defined as a “theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process”.

Inspired in part by VSD, we have developed a process geared towards embedding ethics at the core of CARESSES, an international multidisciplinary project that aims to design the first culturally competent SAR for the care of older adults. Here we describe that process, which included extracting key ethical concepts from relevant ethical guidelines and applying those concepts to scenarios that describe how the CARESSES robot will interact with individuals belonging to different cultures. This approach highlights the ethical implications of the robot’s behavior early in the design process, thus enabling researchers to identify and engage with ethical problems proactively.

I. INTRODUCTION

In recent years, rapidly ageing societies and dramatic shortfalls in the numbers of professional caregivers [1] have translated into increased attention to research on the development of socially assistive robots (SARs) for elder care. This development has prompted a great deal of ethical reflection on the future of SARs in caring roles. Ethical concerns have emerged, for instance, relating to a loss of human contact and dignity, diminished autonomy and privacy [2], and to negative impacts on standards of care [3].

While ethical reflection and analysis have furthered ethicists’ understanding of the interplay between technology, robotics and healthcare, they have often taken place far from the settings where technological innovation is practiced and may be of limited use when considering how ethical considerations can be effectively embedded into the design of future SARs [4].

Different frameworks have been proposed to handle the ethical dimension of technology from within the design and development process, including Value Sensitive Design (VSD) [5]. VSD is rooted in the assumption that design issues in technology are informed by morally relevant considerations, and that such considerations need to be taken into account throughout the design process. It typically consists of conceptual, empirical and technological investigations. The conceptual investigations examine the moral questions associated with the development and use of a given artifact. The empirical investigations, instead, explore the human context in which the artifact will function, while the technical investigations attempt to clarify how existing technologies support or interfere with certain values, or inform the proactive design of technologies prompting the values identified earlier in the VSD process itself. [5]. These investigations are carried out iteratively, giving designers the opportunity to modify design in a continuous manner.

Drawing in part on VSD, we have developed a process geared towards embedding ethics in CARESSES, an international multidisciplinary project that aims to design the first culturally competent SAR for the care of older people [6]. Our process included two main steps: a conceptual phase in which we examined the moral questions linked with the use of assistive technologies with older individuals, and a technical investigation in which we tried to understand how interactions with the CARESSES SAR may hinder or promote the values identified in the first step, in order to proactively identify and address any ethical issues. The conceptual phase was carried out by analyzing relevant ethical guidelines (Alzheimer Europe's Guidelines and Position on the Ethical Use of Assistive Technologies for/by People with Dementia) (hereafter referred to as AEG) to extract key ethical concepts [7], while the technical investigation involved applying those concepts to the scenarios that have been developed within CARESSES to describe how the CARESSES SAR will interact with older adults belonging to different cultures.

The aim of this paper is thus to propose a straightforward approach that highlights, during the design process, the ethical implications of a robot’s behavior and interactions with older adults, and enables researchers to engage with ethical problems early on in the design process rather than restrict ethics to retrospective assessments.

The article is organized as follows: Section II provides an overview of the CARESSES project and briefly describes the CARESSES scenarios and their role within the project; Section III provides background knowledge on the AEG; Section IV describes the analysis we carried out to extract key ethical themes from the AEG, and how those themes were then used to develop an ethically labeled version of the

L. Battistuzzi and A. Sgorbissa are with the University of Genoa, Via all’Opera Pia 13, 16145 Genova, Italy.

C. Papadopoulos is with the University of Bedfordshire, Park Square, Luton LU1 3JU, United Kingdom.

I. Papadopoulos and C. Koulouglioti are with the Middlesex University Higher Education Corporation, The Burroughs, Hendon, London NW4 4BT, United Kingdom.
scenarios; Section V presents three of the ethical themes, including sample extracts from the AEG and from the scenarios, showing where and why the themes are relevant. We close with Conclusions in Section VI.

II. CARESESSES: THE PROJECT AND THE SCENARIOS

CARESESSES is a joint EU-Japan research effort that aims to bring cultural competence to socially assistive robotics. Cultural competence is a concept developed within the nursing domain to describe the capacity to provide effective healthcare that takes into consideration people’s cultural beliefs, behaviors and needs [6]. It will enable the CARESESSES SAR to adapt the way it interacts with older persons to their culture, customs and manners, and be sensitive to their culture-specific needs and preferences [7]. The robot’s capabilities will include:

- communicating through speech and gestures;
- assisting users in daily routines;
- providing health-related assistance, e.g. reminding users to take their medication;
- providing easy access to technology, e.g., internet, video calls, smart appliances for home automation;
- providing entertainment, e.g., reading aloud, playing music and games.

In order to explore the added value of cultural competence, the CARESESSES SAR will be tested in experimental trials with older adults and their caregivers at care homes in Europe and in Japan. Whether the CARESESSES SAR actually achieves the expected levels of cultural competence and the impact that competence has on study participants will then be evaluated by administering standardized questionnaires and interviews to the participants after the trials.

In CARESESSES, 60 different scenarios have been developed by project partners who are experts in Transcultural Nursing. Drawing on the theory and practice of cultural competence in professional caregiving [8], each scenario describes a situation or activity and indicates the robotic capabilities needed to respond to the older person in the specific situation, in a culturally appropriate, sensitive and acceptable way. The scenarios are one of the key drivers of CARESESSES, as they provide the basic groundwork for the development of guidelines for cultural competence in robotics, of the required robotic capabilities and for the design of the experimental trials with older adults and their caregivers. A detailed description of the scenarios and guidelines is available for download from the CARESESSES website [9].

III. THE AEG

The AEG contain general guidelines for ethical decision-making related to the use of assistive technology for older adults with cognitive impairment. They are targeted at a wide, varied audience including caregivers, policymakers, researchers, assistive technology designers, and users, and are based upon literature reviews and consultations with experts and stakeholders.

The authors point out that absolute recommendations based on particular ethical principles are often impractical, if not undesirable, as situations are seldom identical. Therefore, they suggest that the AEG should not be adopted as they are, but require critical appraisal and adaptation. In CARESESSES, this adaptation involved thinking of the AEG in terms of how they could apply to the CARESESSES SAR, and their relevance in a research setting with older adults who have sufficient cognitive competence to participate in the experimental trials.

Along with the guidelines themselves, we decided to include in our analysis Alzheimer Europe’s background document “Ethical issues linked to the use of specific forms of assistive technology” [10], which provides a rich backdrop for contextualizing the guidance provided by the AEG; for brevity’s sake, we will hereafter refer to both documents jointly as the AEG.

IV. EXTRACTING ETHICAL GUIDANCE FROM THE AEG AND LABELING THE SCENARIOS

A. Identifying the ethical themes

To extract the main ethical themes from the AEG we utilized a rigorous, well-established approach termed Qualitative Thematic Analysis (QTA) [11]. QTA is used to identify, analyze and report patterns within qualitative data such as texts and transcriptions of interviews or focus groups, which cannot be numerically expressed. We began by reading through the AEG and identifying the subsections of the document that are not relevant to our project (e.g. where issues having specifically to do with lack of cognitive competence are discussed), which were not considered further. We then analyzed the remaining text, divided it into meaningful segments (paragraphs, sentences or phrases), and, where appropriate, tagged these units of meaning as instances of ethically relevant concepts, in a process termed “coding” the units. While QTA usually proceeds by fragmenting the text into first-level conceptual units and then examining them to establish whether they can be organized under higher-level, more general concepts, we directly aimed for identifying broad ethical themes, which we then used as labels to highlight ethical significance when analyzing the scenarios in the second stage of our approach.

We identified nine broad ethical themes within the AEG: Attachment and authentic interaction; Autonomy; Beneficence, quality of life and well-being; Culturally determined values and preferences; Dignity and personhood; Informed consent and shared decision-making; Preventing harm; Privacy; Stigma.

B. Applying the ethical labels to the scenarios

Having identified the main ethical themes in the AEG, we then used those themes as labels to highlight the ethically relevant elements in the scenario tables that describe the activities to be performed by robots when interacting with older persons during daily routines1. The scenario tables also provide guidance for all the subsequent research and technological development activities, including the

---

1 When this was carried out, 32 scenario tables had been developed involving female users; 28 additional scenarios with male users were added at a later stage.
implementation of motor, perceptual, cognitive and verbal robotic capabilities. Rather than label the tables themselves, which are rich in narrative elements and background information that are not strictly related to the robot, we analyzed the robot tasks that are contained in the scenarios.

Figure 1 shows part of an ethically labeled scenario table. The first line describes what the robot can do in the scenario, the first column includes optimal tasks, the second column lists alternative tasks that take into account current technological limitations, and the third column defines the ethical relevance of the task including reference to the ethical themes identified in the AEG. The second line describes the basic capabilities required along with corresponding software functions which are already available for implementation, and includes the ethical themes legend.

V. ILLUSTRATING THE APPROACH: THREE ETHICAL THEMES IN THE CARESES SCENARIOS

In our approach, the ethically labeled scenario tables are accompanied by an explanatory document. The aim here is not to conduct in-depth analyses but rather to provide researchers who are not ethics experts with straightforward reference and illustrative examples.

Below is an overview of three of the ethical themes that were most frequently identified in the AEG: Attachment and authentic interaction; Autonomy; Privacy. The format is the same as in the actual explanatory document we produced, and is generally the following: first a brief introductory paragraph, then relevant extracts drawn from the AEG (quoted between inverted commas), then extracts drawn from the scenarios (quoted between inverted commas) which have been labeled as associated to the theme in question, and comments where appropriate.

In several instances extracts from the scenarios can be viewed as referring to more than one ethical theme. However, for sake of simplicity and clarity, themes are discussed separately.

A. Attachment and authentic interaction

SARs are often built to resemble humans and to have human-like qualities or behaviors. Because of this, a person may become attached to the robot and even consider it “real” or “alive”. The robot, however, may have to be shared with others – other residents in a care home, for instance – or break down, or not live up to the person’s expectations, leading at times to feelings of jealousy, disillusionment or disappointment [12]. Ethical concerns may also emerge due to the fact that robots may emulate empathy or display caring behaviors; this is especially relevant with older people whose ability to interact with fellow humans may be failing.

Sample extracts – AEG

- “Some people may feel that there is something uncanny about a computer which seems to show concern or exhibit caring behavior, whereas others might appreciate this or simply have no problems with it.”
- “Ethical questions may nevertheless arise such as whether it is right to give SARS of this kind to people whose ability to interact with humans is declining, who might at some point be unable to understand that something is not “alive” and who may be feeling isolated.”

Within the scenarios, the robot tasks that are relevant to this ethical theme are those in which the SAR praises, congratulates, shows interest, comments or simply keeps company with the older person, since all of these actions can be viewed as replicating human behaviors that characterize companionship and caring.

Sample extracts – Scenarios

- “Show interest in Mrs Yamada’s praying customs”
- “Keep company with Mrs Smith while she is eating.”

Although interaction with SARs may involve the risks described above, it can also, as the AEG note, contribute to making a person feel cared for or draw that person into interaction. By expressing interest and care, the SAR can relieve feelings of isolation [13] and, to some extent, meet the older person’s need for companionship [14].

B. Autonomy

Autonomy is often defined as a person’s ability to make her own decisions about her life, based on her principles, values, beliefs, priorities and goals. Traditionally considered to be linked to self-determination, exercising choice, and informed consent [15], it provides one of the core principles of practice with older people.

Sample extracts – AEG

- “SARs could be used to carry out certain tasks on behalf of the user or to provide a sufficient level of support to enable the person to complete the task herself. This should promote the autonomy of the user.”
- “AT can be used as a memory aid to enable people to carry out tasks which would otherwise be difficult or impossible to accomplish on their own due to difficulties remembering what to do, which items are needed for the task, or in which order to carry out each stage of the task.”

Many of the tasks that the robot performs in the scenarios (e.g. bring objects, remind, suggest, store or retrieve information) precisely reflect one or both of the goals that the AEG identify: to promote user autonomy by carrying out certain tasks on the user’s behalf or by behaving as a memory aid.

Sample extracts – Scenarios

- “Ask Mrs Yamada if she wants to make a shopping list and prepare it on the tablet”
- “Ask Mrs Smith when she will see her son again and store the information”

Autonomy has the following sub-themes: independence, proportionality, and balancing autonomy and safety.
1) Independence

According to the AEG:

“People with dementia express satisfaction about not having to rely on other people, so in some cases an advantage to AT could be that it enables people to express their independence.”

SARs, which generally do not give physical aid, may be viewed as supporting and maintaining the independence of older persons because the suggestions, cues, reminders and encouragement they provide help users take care of tasks on their own, care for themselves and manage their life. The same extracts listed above to illustrate how autonomy plays out in the scenarios can be conceptualized as instances in which users’ independence is fostered.

2) Proportionality

Closely linked to the promotion of autonomy and independence is the risk of over-reliance on the robot. According to the AEG:

“It is important to respect the principle of proportionality, which means that the level of intervention should be restricted to what is really needed for a particular person in a particular situation. Providing more assistance than is actually needed may result in the premature loss of capacities which may foster a new form of dependency, namely on AT rather than people.”

This suggests that, should SARs be incorporated in care settings, regular revision of how they interact with users in terms of the type and amount of support they offer will be required in order to ensure it is appropriate for users’ needs, preferences and capacities.

3) Balancing autonomy and safety

Many of the ethical dilemmas that arise with older adults in care settings revolve around how risk should be managed to ensure safety while respecting older persons’ autonomy.

According to the AEG:

“When faced with decisions about the use of AT which necessitate balancing the management of risk with the promotion of the person’s autonomy, the following should be considered: the real rather than hypothetical risks involved; the necessity to focus on risks to the individual and not primarily on risks to the establishment (care home or hospital); considering potential benefits at the same time as potential risks; people have different perceptions of risk and of what level of risk is acceptable to them; it is unrealistic and even undesirable to try to eliminate every possible risk in the life of the person with dementia. Moreover, this would be likely to have a negative impact on their quality of life.”

Analyzing the robot’s tasks in the scenarios reveals significant instances in which it could be used to mitigate risks or ensure user safety. For example:

“Remind Mrs Smith to take her medication if needed”

“Remind Mrs Chaterjee to check that the stove is off”

While such reminders may provide valuable, independence-enhancing solutions to typical problems, we cannot rule out that they could be perceived by users as controlling or interfering with their autonomous decision-making. One way of balancing respect for users’ autonomy with using the SAR to mitigate risks and ensure user safety is to ask for the older person’s consent to such a use in a personalized, shared decision-making approach. For instance, it might be possible to offer the user and caregivers the option to consider a range of reminders that can be activated, and choose different quantitative and qualitative features (e.g. frequency and intensity) for each type of reminder.

C. Privacy

Privacy refers to the ability of individuals or groups to choose which information about themselves they want to share with others. Sharkey and Sharkey have expressed concerns about the potential loss of privacy when SARs are used in the care of older persons [2]: “Older adults might not like to find that an operator could remote control a robot to peer round their apartment before they are dressed, or when they are taking a bath. They might prefer the robot to have to do the equivalent of knocking on the door and waiting to be invited in.”

While they repeatedly acknowledge the importance of respecting older persons’ privacy, the AEG also mention a different view of the relationship between SARs and privacy, pointing out how relying on robot assistance may actually enhance privacy and not interfere with it.

Sample extracts – AEG

“Although SARs are not human, programmed reactions towards the behavior and movement of the user may affect users’ sense of privacy in that people may feel that they are not alone (which can also be positive) or that they are being watched. If the user can control privacy levels, such problems can easily be overcome. SARs may also increase the level of privacy by side-stepping the need for human assistance for tasks which are potentially embarrassing or private.”

In the scenarios, privacy mainly comes to the fore when the user is getting dressed, is using the SAR for video-calls, has visitors, or is engaged in activities such as meditation or prayer. Considerations about privacy, however, are also relevant when the SAR provides assistance by checking emails or messages.

Sample extracts - Scenarios

- “Provide privacy to Mrs Yamada while she is talking to her family”
- “If in the room, provide privacy, observing Mrs Chaterjee quietly during prayer”
- “Inform Mrs Smith if she has any text /telephone” messages and read them to her.”
<table>
<thead>
<tr>
<th>Left: What the robot can do in this scenario.</th>
<th>Middle: Alternative tasks.</th>
<th>Right: Ethical labels.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A6. Prepare a tray with food.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A7. Bring the tray to Mrs C./to the table.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A8. Respond to Mrs. C.’s request to listen to news on the radio.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A9. Keep company with Mrs C. while she is eating.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A10. Inform Mrs. C. of any phone messages and read them to her.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Left: Robot motor capabilities required.</th>
<th>Middle: Corresponding robot API (if any).</th>
<th>Right: Ethical themes legend.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1. Grasp objects (A3).</td>
<td>No dedicated module. Feasible if payload &lt; 300 g.</td>
<td>Ethical themes legend:</td>
</tr>
<tr>
<td>M5. Avoid unexpected static or moving obstacles (A2, A3, A7).</td>
<td></td>
<td>C = Culturally determined values and preferences.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D = Dignity and personhood.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I = Informed consent and shared decision-making.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P = Privacy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PH = Preventing Harm.</td>
</tr>
</tbody>
</table>

**Figure 1.** Scenario table with ethical labels. The first line describes what the robot can do in the scenario, the first column includes optimal tasks, the second column lists alternative tasks that take into account current technological limitations, and the third column defines the ethical relevance of the task with reference to the ethical themes identified in the AEG. The second line describes the basic capabilities required, the corresponding software functions and the ethical themes legend.
As the user may prefer not to share the content of these messages with other people in the room, carrying out this type this type of tasks should be explicitly required or authorized by users.

Respecting the privacy of older persons should go hand-in-hand with respecting the privacy of carers, family and friends.

Sample extracts – AEG

“Video-phoning may be perceived as an invasion of a person’s privacy. Whilst someone is sitting in front of the screen talking to someone, anyone who walks past may be captured on screen against their will or without their knowledge.”

Careful positioning of the SAR in the room and appropriate scheduling of video-calls might help ensure the privacy of all those who are present. Importantly, in CARESSES’ trial, any video data acquired by the SAR will be processed to extract anonymous information and will not be stored, thus respecting the privacy of research participants and of any other person in the room at the time.

VI. CONCLUSIONS

Overall, we added 478 ethical labels to the scenarios and found that, by impacting on users in terms of their autonomy, well-being, privacy, safety and dignity, virtually all of the SAR tasks in the scenarios have ethical relevance.

As the scenarios will provide the basis for the development of guidelines for cultural competence in robotics, of the required robotic capabilities, and for the design of the experimental trial, the ethical dimension of the project that has been brought to the surface can be easily re-examined at each stage. This will give CARESSES partners multiple opportunities to engage with and address any ethical issues, as required to comply with the AEG and with the overall goal of developing practical technological solutions that will promote the autonomy, dignity, privacy and quality of life of older users.

We believe that the strength of our approach is two-fold: first, it shows how VSD-inspired approaches need not rely on a novel examination of the moral questions associated with the development and experimental use of a given artifact, but can incorporate existing ethical guidance; second, it highlights how human-robot interaction scenarios can play a useful role in allowing researchers to identify, engage with and address ethical issues early on in the design process, so that ethical considerations can be effectively embedded into the design. We hope that, with suitable modifications, it may be usefully applied in other research projects involving human-robot interaction scenarios and SARs.

ACKNOWLEDGMENT

This work was supported by the European Commission Horizon2020 Research and Innovation Program under grant agreement No. 737858 (CARESSES).

REFERENCES


