

ESCAPING THE BENEVOLENT ARTIFICIAL PHYSICIAN: PRIORITIZING CARE ETHICS IN AI-BASED HEALTHCARE

Stacy A. Doore, Ph.D. and Jaime Yockey

Department of Computer Science, Colby College, Waterville, ME, USA

sadoore@colby.edu; mcyock@colby.edu

EXTENDED ABSTRACT

Recent reports estimate that 61 million Americans are affected by a chronic health condition that impacts daily life activities ranging from mobility, cognition, hearing, vision, independent living, and self-care (CDC, 2023). Medicare enrollment is expected to double over the next 15 years, leading to more than 80 million beneficiaries by 2030 (U.S. Centers for Medicare & Medicaid Services, 2021). A large percentage of those with chronic healthcare needs are the result of aging and this has produced a significant shortage in skilled, reliable caretakers in residential facilities and home assistance (Fiorini, 2019; U.S. Department of Labor, 2022). Hospital administrators are hoping that AI powered robotics will provide viable solutions to understaffed medical facilities, home care assistance, and those affected with serious healthcare challenges (Bohr, 2020). However, we argue that for emerging technologies such as care robots to be responsibly integrated into current healthcare sector there needs to be a discussion about how to design and implement based on a care ethics (Gilligan, 1982; Noddings, 2013) framework. This paper begins with a brief discussion of van Wynsberghe's care-centered value sensitive design (CCVSD) framework (2013) with a recommendation to include the principles of justice, transparency, and dignity. It uses a fictional narrative to illustrate why there should be a temporal component in the framework to prevent any shift of foundational values in a system designed and deployed in a specific healthcare context.

Background

Within the field of healthcare robotics, there are several distinctions between the types of healthcare settings and the care tasks they perform. Hospital robots serve a similar function as traditional medical assistants whose primary function is to perform non-critical tasks of monitoring or lifting (Kyrarini et al, 2021). Surgical robots help surgeons with fine precision tasks during surgical procedures. Assistive robots are designed to help patients with activities of daily living (ADL) when there are health conditions such as involuntary movement, limited range of motion, and mobility limitations (Yamazaki et al, 2012). Care robots, the focus of this paper, provide nurses with assistance in more complex patient care tasks, collecting vital health metrics and providing social companionship and interactions for vulnerable patients in hospital or rehabilitation settings (Sharkey & Sharkey, 2010; Vallor, 2011). van Wynsberghe (2013) provides a method for classifying care robots based on three dimensions: application domain (healthcare setting), healthcare use (care practice/tasks), and intended users (giver or receiver of care).

In this paper, we use van Wynsberghe's care-centered value sensitive design (CCVSD) framework (2013) to analyze the application of care robots in a fictional narrative to illustrate

the ways in which there are often conflicting values systems at play with the introduction and use of care robots into healthcare settings that changes over time. Van Wynsberghe centers her framework around Tronto's (2010) fundamental care values of *attentiveness*, *responsibility*, *competence*, and *reciprocity* and provides a set of methods to address each one of these concepts during the design phase by examining what these would look like in a specific context with and without the presence of a care robot. While we agree this is a sound place to start, we also believe the framework should be extended to include a temporal component and suggest including additional concepts of related to care ethics such as justice, transparency, and dignity. We now illustrate the rationale for these recommendations based on a narrative about how emerging technologies created through a lens of care ethics and responsible design practices can change over time to produce artificial systems that do not reflect the original values of the designer.

Caring to what end?

In the speculative narrative, *Escaping the Caring Seasons* (Pinkster, 2018), Zora and Anya Stein wrestle with some of humanity's deepest concerns surrounding the future use of AI-based care robots. As a former developer of assistive living facility in a near future setting, Zora designed a rehabilitation hospital that utilized caregiving and diagnostic decision-making AI robots and systems including an AI robot (DOC) to ensure patients were able to return to the comfort of their homes as quickly as possible. As the creator, Zora thought she had embedded a set of values reflecting central premises of care ethics that prioritized a return to independence, relationships with care staff, and communication between systems and staff to build efficiency into an elder care facility. Although a value sensitive design approach (Friedman, 1996) was not explicitly mentioned in the text, the reader is given the impression that this computer scientist was intentional and proactive in the way the system was designed to promote a set of fundamental values based on stakeholder input in the care of patients. However, over time, the hospital that Zora had worked for was acquired by a larger corporation, which made significant changes to the system of care robots to increase automation, reduce on-site administration costs, and maximize profits resulting in a complete loss of autonomy for the patient.

This fictional scenario illustrates the difficulty of maintaining a commitment to an original set of human values in the face of sweeping automation and removal of humans from care roles. Zora is faced with her own loss of autonomy as a caregiver when the imposed restrictions and limitations on Anya prevent her from making decisions for herself and her wife. Seniors in this nursing home are heavily surveilled, their lives are dictated by the care robot's decision-making program through pervasive computer vision sensors and implanted biometric chips that are engaged in the tracking and calculation of their personal health data. Through this system of artificial beneficent care, residents have lost all personal freedom in a sociotechnical system that treats elder care as family burden to be relieved and perhaps even refashioned into a source of popular entertainment. Zora, as the creator of the system of AI powered robot 'caregivers', witnesses the evolution of this system and its shifting of core values and definitions of care over time.

Ethical analysis

Although a technology design and development process may be grounded in a care ethics framework (i.e., CCVSD) that includes all stakeholders to establish its ethical development and use guidelines, the decisions and actions made by humans and AI powered care robots have the potential for harmful consequences for patients and their families. By identifying the moral assumptions in Pinkster's fictional AI care system, we can identify the ways near future systems may fail to account for factors that have significant impacts on an individual's quality of life (or end of life). First, the system has evolved to reduce patient 'wellness' to the state of bodily functions metrics. Second, the lead engineer assumes that embedded values at the design and development phase will remain constant as a system evolves over time. Third, the deployment of care robots for small care tasks to increase human time for meaningful care activities that required the core values of attentiveness, responsibility, competence, and reciprocity (van Wynsberghe, 2013) may lead to the eventual removal of all human caregivers without checks and balances for maintaining its core principles of care ethics.

We suggest there are several other concepts to be considered in all phases of care robotics adoption, beyond the CCVSD four core concepts, that address harms resulting from the erosion of human care giving as illustrated in the example narrative. This move towards efficiency over time is something that van Wynsberghe (2020) also concludes is a potential weakness in her conceptual framework. In response, we propose applying Held's (1995; 2006) "meshing" of justice and care ethics to the CCVSD framework and to stress the need for explicit transparency of decision making by intelligent systems to ensure the fairness in access to emerging care technologies to ensure the system is not based on biased datasets, perpetuating inequities in social systems. In addition, we argue that Ricoer's definition of dignity (1992) that emphasizes community and social relationships in decision making practices should be added to the framework guiding the use of care robots in healthcare settings because ideally the patient should be supported in maintaining their own dignity while those in their social network uphold an attitude of respect to the individual when they are at their most vulnerable (Leget, 2013).

Finally, the addition of a temporal component to the framework moves the values commitment beyond the design and development period to the implementation and auditing stage. The proposal made by Valles-Peris and Domènech (2023), *Caring in the in-between*, calls for practical actions that ensure the consistency of system values over time. This includes the monitoring of relationships and caregiving processes, the engagement of stakeholders to solicit concerns and priorities when making institutional changes, and alleviating fears by instilling freedom of choices in care that are reversible. With these additions, we believe this augmented care ethics framework for the design of emerging healthcare solutions such as care robots may be able to sustain an original set of moral values during the later stages of an intelligent system's deployment and auditing lifecycle.

KEYWORDS: Artificial Intelligence, Care Robots, Ethics of Care, Value Sensitive Design.

REFERENCES

Bohr, A., & Memarzadeh, K. (2020). The rise of artificial intelligence in healthcare applications. In *Artificial Intelligence in Healthcare*. 25-60.

- Bureau of Labor Statistics, U.S. Department of Labor, (2022, September) *Occupational Outlook Handbook*, Home Health and Personal Care Aides, Retrieved June 2, 2023 <https://www.bls.gov/ooh/healthcare/home-health-aides-and-personal-care-aides.htm>
- Center for Disease Control and Prevention. (2023). "Disability Impacts All of Us Infographic." *Center for Disease Control and Prevention*, 15 May 2023, Retrieved June 1, 2023 <https://www.cdc.gov/ncbddd/disabilityandhealth/infographic-disability-impacts-all.html>
- Fiorini, L., De Mul, M., Fabbricotti, I., Limosani, R., Vitanza, A., D'Onofrio, GTsui M., Sancarolo D., Giuliani F., Greco A., Guiot D., Senges E., & Cavallo, F. (2021). Assistive robots to improve the independent living of older persons: results from a needs study. *Disability and Rehabilitation: Assistive Technology*, 16(1), 92-102.
- Friedman, B. (1996). Value-sensitive design. *Interactions*, 3(6), 16-23.
- Gilligan, C. (1982). *In a Different Voice*. Cambridge, MA: Harvard University Press.
- Held, V. (1995). The meshing of care and justice. *Hypatia*, 10(2), 128-132.
- Held, V. (2006). *The ethics of care: Personal, political, and global*. Oxford university press.
- Kyrrarini, M., Lygerakis, F., Rajavenkatanarayanan, A., Sevastopoulos, C., Nambiappan, H. R., Chaitanya, K. K., Babu, A. R., Matthew, J. & Makedon, F. (2021). A survey of robots in healthcare. *Technologies*, 9(1), 8.
- Leget, C. (2013). Analyzing dignity: a perspective from the ethics of care. *Medicine, Health Care and Philosophy*, 16, 945-952.
- Noddings, N. (2013). *Caring: A relational approach to ethics and moral education*. University of California Press. Pinkster, S. *Escaping the caring seasons*. (2018) Roush, W. (Ed.). *Twelve Tomorrows*. 157-179. MIT Press.
- Ricoeur, P. 1992. *Oneself as Another*. London: University of Chicago Press.
- Sharkey, N., & Sharkey, A. (2010). Granny and the robots: Ethical issues in robot care for the elderly. *Ethics and Information Technology*.14, 27-40. Tronto, J. C. (2010). Creating caring institutions: Politics, plurality, and purpose. *Ethics and Social Welfare*, 4(2), 158–171.
- U.S. Centers for Medicare & Medicaid Services. (2021, December). *National Health Expenditure Projections 2021-2030*. Retrieved June 1, 2023 <https://www.cms.gov/files/document/nhe-projections-forecast-summary.pdf>
- Vallès-Peris, N., & Domènech, M. (2023). Caring in the in-between: a proposal to introduce responsible AI and robotics to healthcare. *AI & Society*, 38(4), 1685-1695.
- Vallor, S. (2011). Carebots and caregivers: Sustaining the ethical ideal of care in the 21st century. *Journal of Philosophy and Technology*, 24, 251–268.
- van Wynsberghe, A. (2013). Designing Robots for Care: Care Centered Value-Sensitive Design. *Science and Engineering Ethics*, 19(2).407-433.
- Yamazaki, K., Ueda, R., Nozawa, S., Kojima, M., Okada, K., Matsumoto, K., Ishikawa, M., & Inaba, M. (2012). Home-assistant robot for an aging society. *Proceedings of the IEEE*, 100(8), 2429- 2441.