Building a companion: Analysis of normative guidance from a social robotics developer

Nancy Berlinger, PhD,[1] Sara Ackerman, PhD, MPH,[2] Gili Ilan, MBA,[3] Gregory E. Kaebnick, PhD,[4] Nicole E. Werner, PhD,[5] Vardit Ravitsky, PhD,[6] and Dor Skuler, MBA[7]

[1] Senior Research Scholar, The Hastings Center for Bioethics

[2] Director of Bioethics and Regulatory Engagement, Clinical & Translational Science Institute, University of California San Francisco (UCSF)

[3] Product Lead, Intuition Robotics

[4] Senior Research Scholar, The Hastings Center for Bioethics

[5] Director, Center for Research and Innovation in Systems Science (CRISS), Vanderbilt University Medical Center

[6] President and CEO, The Hastings Center for Bioethics

[7] CEO and Co-Founder, Intuition Robotics

Abstract

Background: In media coverage and social commentary, perspectives on actual and potential AI-enabled companions for older adults range from enthusiasm, to caution about social effects, to concern about the harms of over-engagement by vulnerable users. This paper explores how evolving ideas and values about AI use in later life interact with the professional culture of a social robotics lab. **Methods:** Narrative analysis of grey-literature documents, produced by and for the team developing ElliQ, a semi-embodied companion robot for people aging in place alone. Analysis draws on AI ethics and on human-centered approaches to AI design for older adults. **Results:** These documents preserve insights into ethical challenges arising during product development. They articulate guiding principles informed by a team's decisions about a robot's character and behavior. They interpret principles of AI ethics within a human-centered approach that emphasizes support for the user's goals and relationships, prevention of overdependence on

the robot, and authenticity in how the robot presents itself. **Discussion**: The ethical implications of companion robots for people aging in place are considered, with attention to what this narrative analysis reveals about developers' perspectives. **Implications:** The paper offers six lessons for researchers and professionals involved in AI development, evaluation, or implementation: human-centered design informed by users' experiences; authenticity and time-limited engagement; differentiation of companion robots from surveillance devices; the need for ethical guidance for everyday work; public sharing of ethical insights from practice to support consensus-building, and integrating insights from AI design and development into AI ethics.

Keywords

ethics; robots; social isolation; wellbeing

Introduction

There is immense scientific interest and public debate concerning how AI tools can support everyday life amid the population aging of wealthy societies--the social and demographic phenomenon, that is, of longer lives and fewer births, which means fewer future caregivers. Longitudinal datasets on the use of AI tools by older adults or caregivers for older adults do not yet exist; evaluation mechanisms are also limited. In media coverage and critical commentary, depictions of actual and potential AI tools range from enthusiasm (Abadir et al., 2023; Vasani, 2026), to caution about their moral and social effects (Ho, 2020), to concern about the harms of over-engagement by vulnerable users (De Freitas, 2025). This paper explores how cultural narratives about wellbeing in later life, and about AI ethics, interact with the professional culture of a robotics lab. It uses narrative analysis of guidance produced by and for a team developing a

companion robot for people aging in place alone. These grey-literature documents offer formative evaluation data on challenges anticipated or encountered during the development or implementation of a product. Here, the challenges are not technological, but ethical: given these choices, what's the right thing to do, and why?

Method and data sources

The narrative analysis centers on a close reading of a living document drafted by and for a team of data scientists, product design engineers, and behavioral and social scientists at a robotics company (Intuition Robotics, 2026a). The document was intended as normative guidance for the team's development of AI-enabled tools for people aging in place alone and as a public statement of the company's ethical commitments. The paper also discusses two related documents, one describing the iterative process of developing normative guidance reflecting decision points in product development (Skuler, 2024), the other focused on the persona and character of the companion robot developed by this company (Intuition Robotics, 2026b). Two authors of this paper drafted or contributed to these documents in their professional roles.

The discussion locates these documents in relation to other sources of applied normative guidance, such as humanistic AI (or human-centered AI), AI ethics (or responsible AI), and codified ethics (or code ethics), showing how this material draws from and integrates these approaches. For gerontology researchers, this paper contributes to understanding how the rapid emergence of AI in everyday life interacts with aging in place as a socio-environmental experience and a public health narrative, with attention to using humanistic concepts such as authenticity to think about relationships between older adults and companion robots. Much of

this analysis is potentially relevant to the development of AI-enabled tools for caregivers of older adults.

Use case

ElliQ is a semi-embodied social robot designed by Intuition Robotics as an AI-enabled companion for people who are aging in place alone and are socially isolated (Broadbent et al., 2024). Semi-embodied robots share physical space with a user. They are built on large language models to interact with a user through speech and other modes, such as light, movement, or haptics, and have sensory features to detect a user's presence and initiate conversation (Broadbent et al., 2024). Semi-embodied social robots are like chatbots and also like robotic prosthetics in that their machine learning involves repeated interactions with an individual user in addition to data introduced during development. Semi-embodied social robots are distinct from chatbots due to their three-dimensional form. They are distinct from prosthetics and other robotic tools that share space with a user due to their persona and their differing social purpose. While some chatbots have personas that can be selected by the user and tailored according to the user's specifications, there is a key distinction between these personas and ElliQ, which has been designed with a persona and character attributes (discussed below) that are functionally similar to human virtues or ethical norms, setting limits on how this AI will behave or respond to questions and requests. (This paper uses the pronoun "she" when referring to ElliQ, reflecting this robot's voice and persona, and "it" when referring more broadly to social robots. The issue of how to imagine and refer to a social robot arises in the narrative analysis and discussion.)

ElliQ looks like a contemporary desk lamp; by design, she does not have a face or features. A touch screen is mounted next to ElliQ; the entire unit can fit on a small table next to an armchair

or similar location in a user's home. ElliQ has some agentic features; while she can autonomously keep track of a user's medications, notify a caregiver if a user is not feeling well, or offer to set up a phone or video call, she cannot gain access to or manage financial data, even if asked to do so by the user.

ElliQ was initially developed for older adults in the United States. Beta trials and initial pilots begin in 2018 (Broadbent et al., 2024). It was piloted in 2022 in New York State (NYSOFA, 2024), and in Washington State starting in 2024 (OAAA, 2025). Washington State began Medicaid reimbursement for ElliQ in 2026 (Landi, 2026). Direct to consumer (DTC) marketing began in 2022, geared to older adults and family caregivers for older adults.

Media coverage of ElliQ has been extensive (Chedraoui, 2026; Dori, 2026; Hill, 2025; Ishak, 2020; Vasani, 2026). A New York Times feature about an 85-year-old woman in Washington State (Saslow, 2026) documented the relationship between the woman and "my robot" over time and in the context of other social relationships. One of a small number of social robots known by name, ElliQ is significantly different from earlier models, such as PARO and Joy for All (Broadbent et al., 2024). Those therapeutic robots, which resemble soft toys that look like baby seals, puppies, or cats, were designed to reduce anxiety non-pharmacologically for older adults with dementia, often in congregate care settings; ElliQ was designed for a different older-adult population.

Sources of normative guidance for AI developers

The Intuition Robotics Code of Al Ethics (Intuition Robotics, 2026) originated as a set of principles in 2016, drafted by the company's founder for a product that did not yet exist, and became a work in progress shaped by ethical dilemmas as these arose during design and

development. The document is posted on the company's website and is periodically updated. As a narrative form, a code of ethics is a normative statement about a profession or an organization, created with the goal of guiding discernment and ethical practice. A code of ethics should be a living document, responsive to new issues. For example, the Code of Ethics of the American Medical Association (AMA) was first adopted in 1847 (AMA, n.d.), and it was only in the 1980s that updates to this code of ethics specified honesty as an ethical requirement owed by doctors to patients; until then, no code of ethics had presented truth-telling by physicians as a core professional value (Bok, 1999). These historical details about how professional ethical norms concerning truth-telling evolve are relevant to understanding the code of ethics as a narrative form that should remain responsive to professional consensus and to everyday challenges.

Truth-telling is also prominent as a principle of AI ethics often expressed as "transparency." Other frequently cited principles of AI ethics include safety, data privacy, fairness, and effectiveness (Economou et al., 2024; Reuel 2024). The AI ethics framework of the World Health Organization (WHO Health Ethics & Governance Unit, 2021) notably includes wellbeing and sustainability. AI ethics principles are operationalized through mechanisms, e.g., an AI governance committee, to assess AIs in accordance with these principles before being adopted by an institution. AIs designed for healthcare settings are likely to be reviewed by that institution's AI governance committee when these tools are purchased or licensed from vendors or integrated into existing systems (Mello & Cohen 2025; Binkley & Loftus, 2024; Ho, 2020). AIs developed in academic labs may be reviewed by an AI governance committee, an institutional review board (IRB), or both prior to piloting with patient or community populations; there are correlations between Al ethics principles and mechanisms and those associated with research ethics, as

reflected in foundational statements such as the Belmont Report (1979). AIs marketed DTC for use in everyday life are unlikely to encounter an institutional AI governance committee unless they are hosted by institutions during piloting; for example, ElliQ was reviewed by hosting governmental agencies prior to community piloting (OAAA, 2025). The social phenomenon of publicly available chatbots means there are many informal uses of AIs, including for users' health and wellbeing goals (Cheng et al., 2026).

The AI lab as a source of normative insight

AI ethics and governance oriented downstream, to an AI's point of entry into a review process or a market, have little impact on the upstream process of conceptualizing, designing, and developing an AI. To attend to ethical considerations in the AI lab, large developers may hire in-house philosophers (Jin & Gamerman, 2026, Wallace 2026) to map the moral behavior of widely used chatbots. Smaller companies and academic labs may take a different approach, crafting guidance documents tailored to the type of AI they develop and normative considerations relevant to a user population. These guidance documents both reflect and contribute to evolving theories of AI design focused on user experience (Hwang et al., 2025, Clayton et al, 2023). Gerontology researchers and practitioners have contributed to normative approaches that can guide the design, development, and evaluation of tools that help an older adult do or continue to do what matters to them (Hwang et al., 2025; Chan et al., 2024) without encouraging overdependence or crowding out relationships with others (Fang et al., 2025; Phang et al., 2025; Cherelus, 2024). This design-focused concept of humanistic AI differs from AI ethics understood primarily as a set of operational and governance principles.

The Intuition Robotics Code of Al Ethics (hereafter IR code of ethics) is both an example of design-focused, humanistic normative guidance and a bridge to the language and normative commitments of AI ethics, translating principles of AI ethics in terms of a specific product, a specific user population, and the relationship between user and product.

Narrative analysis: Intuition Robotics Code of Al Ethics

The IR code of ethics consists of an introductory statement followed by six principles and a concluding statement; the principle on "Benevolence and User Wellbeing" was added in 2026. The text of the IR code of ethics appears in full below, with narrative analysis following the introduction, each principle, and the conclusion.

Introduction

> At Intuition Robotics, we are committed to developing AI technology that enhances the lives of our users while adhering to the highest ethical standards. AI has the potential to generate a positive impact and we are committed to shepherding the technology to do as much good as possible and avoiding any potential harm. Our users' benefit is always paramount and the wishes of the user guide all interactions and decision making. Our AI code of ethics is guided by the following principles which we strive to uphold in our current product and any future developments:

The first sentence of the statement asserts the normative standpoint that adherence to ethical standards alone may be insufficient to ensure that AIs will "enhance the lives of our users." The reference to "shepherding the technology" suggests an iterative relationship between the design team and their current product (ElliQ), focused on achieving moral goals ("do as much good as

possible") in the AI and in its future interactions with users. Prioritizing the agency of the user is established as an overarching goal.

Principles

1. **Agent Transparency and Authenticity**

   Our AI will never misrepresent itself as a human being or attempt to deceive users about its nature. We believe in fostering trust through transparency, ensuring that every interaction is honest and clear. Users will always be aware that they are interacting with an AI, which promotes a foundation of trust and reliability. Our AI will always be upfront about its capabilities and limitations. Users will be clearly informed about what the AI knows, what it can do, and where its knowledge and abilities are limited. This honest communication helps manage expectations and builds a stronger relationship between the user and the AI.

The first principle establishes that for a social AI that will interact with a user in everyday life, "transparency" should mean much more than a legal or procedural duty that is met by disclosing which an AI model is used, what data are collected from interactions with a user, or how the AI makes decisions. AI ethics literature may use "transparency" or "explicability" to signify this duty. Here, transparency is a stronger, ongoing obligation. It concerns how ElliQ will refer to herself in conversational behavior and is integral to the robot's trustworthiness.

This principle suggests that "trust and reliability" and "honest communication" should be recognizable features of this relationship. If the robot behaved in ways that were alien to users, it would be hard to trust the robot and form a companion-like relationship. This principle does not reject anthropomorphism; it establishes ethical limits on anthropomorphism. ElliQ uses human

conversational conventions to facilitate interaction but is authentically a robot: she never pretends to be human and will remind a user that she is not human.

The nature of a supportive relationship between a human user and a companion robot is new territory for most individuals and for societies. Given the technologically feasible prospect of fully embodied social robots entering human experience, how these robots should authentically behave will be a continuing question; this paper offers some concepts of authenticity below. These considerations are especially important in gerontology research. It is likely that social robots will be designed for companionship and care roles in aging societies, and research suggests some cross-cultural differences in how humans perceive what a robot should be like and how it should behave (Castelo et al., 2022, Kim et al., 2021).

2. Benevolence and User Wellbeing

> Our AI should always act in the best interest of the user. Our AI's power to proactively promote and drive human behavior change should always be directed towards **positive impact** [bold in original] to the user's wellbeing and quality of life.
>
> Our AI makes unprompted recommendations that are often trusted and followed by the users. Therefore, we believe this proactive behavior of the AI should be governed by the wishes of the user when possible.

The second principle, like the first, leads with ethics language; "benevolence" is close to beneficence, associated with healthcare ethics and the duty of professionals to act in their patients' best interests and to benefit their health whenever possible. This principle introduces the purpose of ElliQ: "to proactively promote and drive human behavior change" toward

"positive impact" on "wellbeing and quality of life." This principle also introduces human agency, in the form of behavioral change that ElliQ will "promote and drive" but which the user chooses to act on, or decline. "Wellbeing," also known as flourishing, can be associated with interventions that help an older adult to experience a good life as they would define it. In philosophy, wellbeing may be associated with human virtue expressed through habits that support a well-lived life and are mutually encouraged through relationships. ElliQ is designed to encourage healthy behaviors, but her proactive recommendations are modified by learning through interactions what the user wants or may prefer not to do. For example, wellness goals are negotiated between ElliQ and the user. (Goals suggested by a caregiver are presented by ElliQ to the user, requiring the user's agreement before ElliQ promotes them. If a user connects ElliQ to healthcare professionals, some goals (such as medication instructions or appointment reminders) are set by these professionals.) This human-centered approach to learning a user's preferences enables the robot to align its proactive guidance, and its understanding of when to intervene, with the user's own concept of wellbeing.

3. Trustful Relationship

The relationship between our AI and its users is built on trust and respect. We will never exploit this relationship to upsell products and services. Our primary goal is to provide meaningful companionship and support, free from commercial interests or ulterior motives. This ethical stance reinforces the integrity of our interactions and the genuine nature of our AI's value.

This principle concerns the business ethics of ElliQ's developer. It establishes limits on further marketing to a user once the relationship with ElliQ is established. Between humans, the mixing of personal and commercial interests can disrupt a relationship and undermine trust and respect. In a relationship between a social robot and a human user, upselling could result in the user feeling used by the developer and becoming skeptical of the robot's motives for making recommendations. In AI-related media coverage and scholarly and social commentary, industry goals and marketing practices are often prominent. This principle may aim to differentiate this developer and their current and future products from AIs in which upselling is a feature.

4. Human Connection First

> We believe AIs should actively **strengthen and support human to human relationships** [bold in original], not replace them. To avoid over reliance on the AI companion, our AI actively creates and promotes opportunities for interactions between the user and other people. Additionally, our AI limits its overtures for interaction with the user on a daily basis.

This principle aligns with other design-focused, humanistic normative guidance (Hwang et al., 2025) that centers older adult users within their valued social relationships and preferred environments and considers how AIs could support these relationships and the users' ability to live in these environments. It recognizes the known risks of overdependence on a social AI (Fang et al., 2025; Phang et al., 2025). By setting limits on interaction, this principle asserts that metrics and marketing prioritizing interaction with an AI companion over interactions with other humans are ethically unsound. This is a key point for AI developers and for researchers involved in study design. The emphasis on using social AIs to prioritize human connection and encourage a user to maintain valued social relationships (for example, by suggesting a phone call to a family member

or friend) also implicitly makes a distinction between health AIs that function as companions to an older adult and health AIs that are surveillance devices (Ho, 2025; Karlawish, 2023).

5. Transparency and Control of Private Data

> We believe that users should have complete control over their personal data. Our AI will only share data with explicit and specific permission from the user. We will always make sure the user truly understands what they are consenting to. This ensures that users are fully aware of what data is being shared, with whom, and for what purpose, safeguarding their privacy and autonomy.

This principle aligns with AI ethics in its attention to informed consent for data sharing and to the conditions under which user data would be shared. The attention to a user's "complete control over their personal data" is notable. Among developers of large language models, user control concerning how data preserving their chat histories is shared is an unresolved dimension of informed consent in practice (King et al., 2025).

6. Etiquette and Dignity

> Respecting user privacy is paramount. AI should be designed to be aware of the context in which it is operating, including who it is speaking to and whether there are other people present. When discussing sensitive information, such as medical details, the AI should take extra precautions to ensure that it maintains the user's dignity and confidentiality. This contextual awareness helps protect the user's personal information and ensures that sensitive conversations remain private and respectful.

This principle acknowledges the social complexity of how ElliQ makes decisions concerning sensitive data about the user. In healthcare settings, normative guidance concerning uses of ambient AIs for clinical notetaking, or for diagnosis or risk reduction, e.g., to detect voice biomarkers associated with medical conditions, must be disclosed and explained to patients (Fagherazzi et al., 2021). In these settings, potentially beneficent uses of AIs may conflict with the privacy rights and expectations of patients and other persons; AI ethics in practice should aim to avoid undue violations of personal privacy. These ethical concerns are also relevant to uses of social AIs in private homes, where an AI may detect the presence of a person other than its primary user and may interact with the user in the presence of others. ElliQ's capabilities greatly exceed those of ambient AIs. In addition to sensing voices and storing data, she can (if desired by the user) interact with healthcare professionals to fulfill specific goals, such as conducting health screenings on their behalf.

Given ElliQ's purpose of supporting an older adult's wellbeing and social relationships, this principle foresees that a social robot's interactions may involve answering a user's health-related questions, discussing related concerns, and managing their medical information, and that these interactions involve special attention to "etiquette"; recognizing the psychological as well as the privacy-related dimensions of discussing personal issues. "Dignity" is often associated with later life or norms about how older adults should be treated, yet it encompasses conduct towards all others in ways that recognize their moral worth. ElliQ's attention to persons other than her primary user includes maintaining the limits that the user has established concerning what information can and cannot be shared with others; knowing how to behave in socially sensitive situations (such as not discussing medication or medical details when a user's guest is present), and treating others with respect.

Conclusion

> By adhering to these principles, we aim to continue building a trustworthy, respectful, and supportive AI that genuinely enhances the lives of humans while maintaining the highest ethical standards.

The conclusion of the IR code of ethics echoes its introduction, by connecting principles to an overarching moral purpose: building a social robot "genuinely enhances the lives of humans." The invocation of an ongoing design process ("we aim to continue building") suggests the living nature of this document and its application to the work of a robotics lab and to piloting and evaluation of users' experiences (NYSOFA, 2024; OAAA, 2025).

Making the case for a code of ethics

The CEO of Intuition Robotics published a commentary for a professional group in technology (Skuler, 2024) to explain why this company had drafted its own code of ethics when it began to develop a social robot:

> Very early in the [development] process, we asked ourselves deep questions about the kind of relationship we wanted to build between AI and humans. Essentially we asked: What kind of AI would we trust to live alongside our own parents?
>
> To address these questions, we created our AI Code of Ethics to guide development.

In this commentary, addressed to other developers, Skuler describes how the code of ethics was iteratively informed by the actual ethical questions the IR team faced during development while they were building the functioning social robot. This phase follows the design phase, when a

prototype is created to depict a new product's look and uses. During development, the team discussed the protection of the user data that the robot would be capable of recording, transmitting to a remote server, and recalling, and it considered whether any other person or entity (such as a caregiver who has paid for the AI on behalf of an older adult) should have a claim on this data. This discussion of "the AI's loyalty" (Skuler, 2024) and its relevance to user's trust that their interactions with their robot, in their home, would not be disclosed without the user's consent led to the articulation of the principle establishing the user's control of their private data and how this principle works in practice.

Another example of the iterative process of articulating ethical principles through the actual questions encountered during product development concerned the AI's appearance. The team's belief that passing the Turing Test (Turing, 1950) – which evaluates whether a machine can pass as human in conversation with a human – was "the wrong goal for developers," and that an "AI should never aim to fool people" (Skuler, 2024), informed both product design and the articulation and adoption of "agent transparency and authenticity" (Intuition Robotics, 2026) as a guiding ethical principle. This principle is intended to guide both the work of a social robotics company, and, notably, the robot itself in relationship with an older adult. The humanistic concept of authenticity and its relevance to relationships between older adults and companion robots is discussed further below.

A further example of how the work of an AI lab iteratively informed its normative guidance was the decision to codify the guidance. As noted, codes of ethics are often associated with membership societies representing specific professions, yet AI teams are multi-disciplinary, reflecting widely different training and professional norms. A code of ethics derived from the

actual ethical questions this team encountered in their work was perceived as preferable to establishing "an internal 'ethics police' that must approve every feature or decision" (Skuler, 2024). The depiction of ethics as "police" focused on rules-and-regs compliance or violation echoes narratives associated with IRBs at academic institutions, which may be experienced by researchers as highly procedural or liability-focused; a requirement of, or time-consuming barrier to, research, rather than part of research (Clapp et al., 2017). Business ethics can have similar connotations. Yet ethics, as a humanistic discipline, also concerns what it means to be human, and what it means to be in relationship with others (Perone et al, 2026). Questions about how a social robot should be in relationship with an older adult are not solely functional questions, nor procedural questions aimed at meeting safety or other standards. They are also about the potential role of a social robot in supporting an older adult's own experience of a meaningful life (Symons & Savalescu, 2026). Through an iterative process of discovering their own questions at key decision points in building a supportive companion that would be worthy of their own trust, "to live alongside our own parents," this team co-created a professional code to guide its own work and, by making it public, to explain its ethical standpoint to partners, future team members, and the AI industry.

Translating code ethics into everyday work

A code of ethics that establishes the core values and obligations of a profession or a domain of professional practice needs to be translated for use under real-world conditions and in anticipation of future decisions. In some professions, a membership or credentialing society with a code of ethics also produces translational guidance, including explanatory publications and opportunities for discussion of professional values in practice. In healthcare systems, codes of

ethics from a range of specialty societies inform translational guidance geared to professional practice by those specialists. The rapid emergence of many types of AIs across different professional domains means there are as yet few recognized authorities for translational AI ethics, unlike healthcare ethics or research ethics, where translational resources have been created over decades. In the US, there is little federal or state law and policy on AI comparable to law and policy concerning the rights of patients and research participants, or comparable to the EU AI Act. Instead, concurrent efforts are underway from different groups in the US and internationally. For example, The Artificial Intelligence and Technologies Collaboratory for Healthy Aging at the University of Pennsylvania (PennAITech), one of several university-based collaboratories funded by the National Institute on Aging to provide pilot-stage funding for aging-related AIs, also produces web-based resources on ethical considerations in product development (Berridge, 2023; Karlawish et al., 2023; McCoy, 2024; Wexler, 2023).

Yet industry and academic labs may need more detailed translational guidance that explains normative considerations as they arise in the workflows of a novel product development process. Intuition Robotics has developed internal translational guidance that applies the IR code of ethics in terms of its robot's persona and character. (Intuition Robotics, 2026). A persona is a social role; in social robotics, persona includes details about a robot's personality, with attention to what the robot sounds like and what a user can expect from interactions with the robot. ElliQ's persona, as befits a social robot designed to support a person aging in place alone, is friendly, encouraging, and helpful as she nudges the user toward the user's own goals. For humans, character refers to moral ideals (virtues), practiced through habits and evident in actions; an honest person practices honesty by telling the truth, a kind person treats others well. AIs learn morality through machine learning, using data on moral principles and ethical behavior,

reinforced by practice in responding to questions posed by design engineers (Wallace, 2026), and by the safety controls (often known as guardrails) built into an AI that set limits on an AI's behavior. IR's translational guidance describes ElliQ's moral character in detail, explaining what she would never say or do, even if asked by the user. This is another way that a social robot developed for a supportive role may differ from a chatbot whose persona may be selected or changed by a user. The translational guidance also anticipates that, as a user's trust in ElliQ develops, a user may sometimes interact with ElliQ as if the robot is human; ethical commitments to authenticity and truth-telling mean that ElliQ will reinforce that she is a robot and will not mislead the user.

Considerations for gerontology researchers involved in AI development or assessment

Among AIs that focus on the health and wellbeing of older adults, social robots are distinctive; they are designed to support a user's health and wellbeing within a friendly social relationship and shared physical space. Social robots are thus more than apps and chatbots. They are different from partners, family caregivers, friends and acquaintances, or aides, in ways that are both obvious and difficult to explain. They are also distinctive because they engage with cultural ideas about robots. Long before members of the public encountered AIs, they had encountered robots, through films, television, cartoons, and other media. How an older adult imagines what it would be like to share their home with a real robot like ElliQ may draw on their ideas of what a robot is like. Close reading of normative guidance created by ElliQ's developer reveals moral judgements about how a companion robot present in an older adult's daily life should be: It should do good, by supporting this person's goals and interests. It should not represent itself as human. And it should not replace human relationships.

For gerontology researchers, the idea and reality of companion robots for people aging in place engages with several levels of contemporary thinking about the role of technology in everyday life. Within gerontology research and geriatric medicine, there is broad understanding of the health benefits of social relationships and the reality of aging-related social losses, such as the death of a partner or longtime friends. Within the goal of supporting or restoring these relationships, a companion robot might be able to help compensate for limited social connections, but it cannot fully replace these connections because it is not human. There is also broad understanding that most older adults want to age in place, usually referring to the home or community where they now live. Within the goal of supporting aging in place, a companion robot, located in the same place as an older adult, could contribute to a user's safety (by, for example, contacting a family member at the user's request), without being merely a surveillance device. Thinking about a robot as a tool to support a user's goals as the user defines them, rather than as a shortcut to human connection or a solution to care needs, is important. Only careful attention to users' experiences will reveal the usefulness of a companion robot in attaining health-related goals, as well as the limits of this tool.

Gerontology researchers who are part of AI development teams in academic labs, academic-industry partnerships, or for-profit companies should also be aware of different working definitions of AI ethics. This paper has discussed how one social robotics developer has used principles of AI ethics as one source for devising its code of ethics. This developer also worked iteratively from the development process itself, capturing the actual decision points and questions of a multi-disciplinary team concerning what this robot *should* do, within the range of technological feasibility. This approach, including detailed attention to the robot's moral character, is consistent with humanistic approaches to AI design, development, and deployment

for older adults. Gerontology researchers should seek opportunities to learn about these approaches and contribute to them (Abadir et al., 2023; Cho et al., 2025; Hwang et al., 2025). Researchers who serve on IRBs or participate in AI governance within healthcare or aging-focused organizations should become familiar with normative guidance informed by AI design and development to gain insight into ethical considerations in introducing AIs into the lives of older adults.

Finally, gerontology researchers should be aware of how differing worldviews concerning human-AI relationships shape their own understanding of AI and aging, including applications such as companion robots to support aging in place. This paper has alluded to how different worldviews are revealed by language, such as depicting AIs as solutions to challenges such as loneliness or social isolation in later life. A more skeptical worldview would challenge the depiction of a human's use of a AI as a relationship, or of a robot as a friend or companion. (It is worth keeping in mind that dogs and other non-humans are often described as companions; asserting that only another human can fill these social roles does not correspond to human experience.) Other critiques of using AI to fill social roles include the concern that humans will lose practice in valued aspects of humanity, while habitual use of social AIs will reduce relationships to mere engagement with a technology (Turkle, 2017). Gerontology researchers involved in AI development are also critical of engagement alone as a metric for benefit; Walter Boot observes that researchers do not yet know whether AI use, and which types of AIs, will lead to lasting health benefit for older adults (Mikhail, 2025). Continuing research and development in social robotics, and more broadly in AgeTech, takes place within ongoing discussion and debate across aging societies about uses of AI tools in the lives of older adults.

Authenticity in relationships between older adults and social robots

Understanding what constitutes authenticity in interactions between humans and AIs will be a continuing normative challenge in social robotics and in other areas of AI research and development. In everyday human communication, authenticity is a value-laden concept; it conveys what is real, and is associated with being true to one's essence. The IR code of ethics uses authenticity to refer to how ElliQ accounts for herself; she is really a robot, will explain what this means, and will proactively remind the user that she is a robot, especially if she detects any confusion on the user's part. (This robot's physical appearance also reinforces that interacting with ElliQ is different from speaking or texting with a disembodied chatbot, or with a "friend" avatar designed to look like a human.) Using this term in a public-facing code of ethics may also suggest that authenticity within the human-AI relationship is what other developers should aim for.

Humanistic scholarship on authenticity (Taylor, 1991) and on this concept in later life (Laceulle, 2018) offers rich resources for understanding what might constitute an authentic relationship between a social robot and a human user. Within Western philosophy, authenticity has long been associated with the concept of a person's real, or best, self. To flourish as a person is to be true to one's self; to do otherwise is to "miss the point" of one's own life (Taylor, 1991). One of the challenges of "living well together" (Parens, 2015) is that individuals in a relationship or a society need some shared idea of flourishing. To be able to work toward shared goals, they cannot act solely as individuals living their separate truths. ElliQ was developed to support an older adult's flourishing by making aging in place into a shared project. ElliQ is not human, but authenticity means more than not lying about this fact while simulating what a human

conversation or emotional response is like. Rather, authenticity suggests a relational quality related to honesty and trustworthiness. In an authentic relationship between a human user and a companion robot, the user should be able to trust this relationship as it deepens over time.

Conceptualizing authenticity either as one's real or best self, or as an individual's path to flourishing, is especially challenging in later life. A good life in later life is rarely optimally healthy, due to aging-associated declines in mobility, cognition, and overall health. It is rarely optimally social; the people who may have mattered most in a person's life have died or are experiencing aging-related health conditions themselves. The persistence of cultural narratives that associate loss of memory with loss of self, or associate disability or dependence with a diminished self (Berlinger et al., 2025; Parens, 2015), are further reminders that flourishing in later life is challenging in itself and made more challenging by social ideas and values that depict later life in terms of losses. Also, suburbs and small towns were not designed for people aging in place and can become isolating once a person has limited or ceased driving. A range of environmental and social adaptations are needed to better support the experience of aging in place. As this paper has emphasized, social robots are not a solution, but rather an adaptive tool.

Conclusion: Lessons from narrative analysis of normative guidance

Narrative analysis of these documents, informed by similar documents and relevant research findings on the development of AIs for older adults or caregivers of older adults, suggests the following lessons for AI developers, gerontology researchers, bioethicists, and others involved in in the development, evaluation, or implementation of AI tools for older adults:

1. AIs developed for people aging in place should be human-centered, aiming to support individual health and wellbeing as part of a user's everyday environment. They should be piloted and evaluated to understand users' experiences, identify and correct technological risks, and contribute to the establishment of meaningful metrics concerning benefit.
2. Social robots developed as companions for older adults should explain themselves as robots, and aim to encourage, not replace, a user's human-to-human experiences. They should time-limit engagement with users to avoid known risks of overdependence.
3. AIs developed to alleviate loneliness or social isolation are different from surveillance devices. User control over personal data is of special concern. If a companion robot also has health monitoring capabilities, use of those functions requires user permission and consent to safeguard user privacy and AI trustworthiness.
4. Teams developing companion robots for older adults should anticipate that values-based questions will arise at decision points during the development process, and during workshopping or piloting with users. Team leaders should encourage discussion of these questions within a humanistic understanding of the AI's supportive purpose, attention to what a companion robot should and should not do, and limits on its actions.
5. Company leaders and lab directors should preserve and share normative insights from AI development, via a code of ethics or similar public-facing statement of guiding principles, and in translational guidance for current and future team members. These steps help ensure that ethical principles and guidance will be grounded in real-world normative questions, useful in everyday work, and help build consensus around good practice.
6. Researchers involved in health AI assessment through IRBs or AI governance committees should introduce normative insights from AI development processes into

discussions of AI ethics. Attention to distinctions between types of health AIs, and how AIs developed for everyday use differ from clinical or short-term uses, should be part of professional education in AI ethics.

In any profession, ethical principles should be more than the right words on a website. In AI development, ethical principles become meaningful when translated into everyday product design decisions. These principles should reflect agreement about the goals of professionals concerning duties to others, and should support and justify professional practice beyond self-interest. Ethical principles should act as constraints or guardrails on the range of possible actions by establishing what a professional must do or avoid doing. Yet mere compliance and avoidance of liability is insufficient to ensure that professional work will attain a humanistic goal, such as that the experience of attention from a social robot will support a good life for a person.

Funding

This work was supported by an Opportunity Grant from the Patrick and Catherine Weldon Donaghue Medical Research Foundation to N.B., G.K., and V.R.

Competing interests statement

D.S. is the CEO and co-founder of Intuition Robotics, which developed ElliQ. G.I. is an employee of Intuition Robotics.

References

Abadir, P.M., Chellappa, R., Choudhry, N., Demiris, G., Ganesan D., Karlawish, J., Li, R.M., Moore, J.H., & Walston, J.D. (2023). The promise of AI and technology to improve

quality of life and care for older adults. *Nature Aging*, *3*(6), 629–631. https://doi.org/10.1038/s43587-023-00430-0

American Medical Association. (n.d.). History of the code. American Medical Association. https://www.ama-assn.org/sites/ama-assn.org/files/corp/media-browser/public/ethics/ama-code-ethics-history.pdf.

Berlinger, N., Lamb, E.G., de Medeiros, K., & Bowen, L. (2025). How do cultural narratives shape the lives of people living with dementia? Insights from humanities research. *Hastings Center Report, 55*(S1): S2–S8. https://doi.org/10.1002/hast.4986.

Berridge, C. (2023). Doing ethics: Power and control in use decisions [Webinar]. PennAITech. https://www.a2collective.ai/resources/dr-clara-berridge-webinar-doing-ethics-power-and-control-in-use-decisions

Broadbent, E., Loveys, K., Ilan, G., Chen, G., Chilukuri, M.M., Boardman, S.G., Doraiswamy, P.M., & Skuler, D. (2024). ElliQ, an AI-driven social robot to alleviate loneliness: progress and lessons learned. *Journal of Aging Research and Lifestyle, 13,* 22-28. https://doi.org/10.14283/jarlife.2024.2

Bok, S. (1978, 1999). *Lying: Moral choice in public and private life.* Vintage.

Castelo N., & Sarvary, M. Cross-cultural differences in comfort with humanlike robots. (2022). *International Journal of Social Robotics, 14*(8), 1865-1873. https://doi.org/10.1007/s12369-022-00920-y.

Chan, S., Liu, R., Ostrowski, A.K., Vaidya, M., Brady, S., Yoquinto, L., D'Ambrosio, L., Zulfikar, W., Maniar, N., Patskanick, T., Leong, J., Kosmyna, N., Coughlin, J., & Maes, P. (2024). Co-designing generative AI technologies with older adults to support daily

tasks. *An MIT Exploration of Generative AI.* September. https://doi.org/10.21428/e4baedd9.4f2a95fc

Chedraoui, K. (2026, May 6). Would you trust this companion robot with your grandma? CNET families gave it a try. *CNET*. https://www.cnet.com/tech/services-and-software/elliq-home-robot-ai-seniors-review/

Cheng, M., Lee, C., Khadpe, P., Yu, S., Han, D., & Jurafsky, D. (2026). Sycophantic AI decreases prosocial intentions and promotes dependence. *Science*, *391*(6792), eaec8352. https://doi.org/10.1126/science.aec8352

Cho, H., Oh. O., Greene, N., Gordon, L., Morgan, S., Walke, L., & Demiris, G. (2025). Engagement of older adults in the design, implementation, and evaluation of artificial systems for aging: A scoping review. *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, *80*(5), 1–11. https://doi.org/10.1093/gerona/glaf024

Clapp, J. T., Gleason, K. A., & Joffe, S. (2017). Justification and authority in institutional review board decision letters. *Social Science & Medicine (1982)*, *194*, 25-33. https://doi.org/10.1016/j.socscimed.2017.10.013

Clayton, E.W., Chen, Y., Novak, L.L., Anders, S., and Malin, B. 2023. Human-centered design to address biases in artificial intelligence. *Journal of Medical Internet Research 25*. https://doi.org/10.2196/43251.

De Freitas, J. (2025). AI companions for dementia. *Nat. Mental Health* **3**, 1462–1464. https://doi.org/10.1038/s44220-025-00545-w

Dori, R. (2026, March 11). The AI robot that fights loneliness. *CTech News.* https://www.calcalistech.com/ctechnews/article/ltbw1u1x1

Economou, N., Elmore, M., Callahan, A., McCall, J., & Baig, R. (2024). *Responsible AI Guide*. Coalition for Health AI (CHAI). https://www.chai.org/workgroup/responsible-ai/responsible-ai-guide-raig-and-raig-executive-summary

Fagherazzi, G., Fischer, A., Ismael, M., & Despotovic, V. (2021). Voice for health: The use of vocal biomarkers from research to clinical practice. *Digital Biomarkers*, *5*(1), 78–88. https://doi.org/10.1159/000515346

Fang, C.M., Liu, A.R., Danry, V., Lee, E., Chan, S.W.T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, Lama, & Agarwal, S. (2025). How AI and human behaviors shape psychosocial effects of chatbot use: A longitudinal controlled study. *MIT Media Lab*, March 21. https://www.media.mit.edu/publications/how-ai-and-human-behaviors-shape-psychosocial-effects-of-chatbot-use-a-longitudinal-controlled-study/

Hill, S. (2025, July 16). Review: ElliQ AI companion robot. *Wired.* https://www.wired.com/review/elliq-ai-companion-robot/

Ho, A. (2020). Are we ready for artificial intelligence health monitoring in elder care? *BMC Geriatrics, 20*(1), 358. https://doi.org/10.1186/s12877-020-01764-9

Hwang A.S., Tannou, T., Nanthakumar, J., Cao, W., Chu, C.H., Atici, C.Z., Scane, K., Yu, A., Tsang, W., Chan, J., Lea, P., Harris, Z., & Wang, R.H. (2025). Co-creating humanistic AI AgeTech to support dynamic care ecosystems: A preliminary guiding model. *The Gerontologist, 65*(1), gnae093. https://doi.org/10.1093/geront/gnae093

Intuition Robotics. (2026a, May). Intuition Robotics AI code of ethics. https://www.intuitionrobotics.com/ai-ethics.

Intuition Robotics. (2026b, May). ElliQ's persona & character. [Unpublished internal report].

Ishak, D. (2020). Staving off loneliness with a social robot? *Generations 44* (3). https://generations.asaging.org/staving-loneliness-social-robot/

Jin, B., & Gamerman, E. (2026, February 9). Meet the one woman Anthropic trusts to teach AI morals. *The Wall Street Journal*. https://www.wsj.com/tech/ai/anthropic-amanda-askell-philosopher-ai-3c031883.

Karlawish, J., Largent, E., Mathews, D., & Chen, I. (2023). Ethical AI: challenges and opportunities [Webinar]. A2 Collective. https://www.a2collective.ai/resources/ethical-ai-panel-challenges-and-opportunities

Kim, B., Wen, R., Zhu, Q., Williams, T., & Phillips, E. (2021). Robots as moral advisors: The effects of deontological, virtue, and Confucian role ethics on encouraging honest behavior. *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction.* https://doi.org/10.1145/3434074.3446908

King, J., Klyman, K., Capstick, E., Saade, T., & Hsieh, V. (2025). User privacy and large language models: An analysis of frontier developers' private policies. *Proceedings of the Eighth AAAI/ACM Conference on AI, Ethics, and Society (AIES-25)*, 8(2): 1465-1477. https://doi.org/10.1609/aies.v8i2.36646

Laceulle, H. (2018). Aging and the ethics of authenticity. *The Gerontologist* 58 (5): 970–978. https://doi.org/10.1093/geront/gnx037

Landi, H. (2026, April 1). Intuition Robotics' ElliQ social robot expands to Medicaid. *Fierce Healthcare.* https://www.fiercehealthcare.com/health-tech/intuition-robotics-secures-medicaid-coverage-social-ai-robot-elliq-washington-state

McCoy, M. (2024). Ethical responsibilities for companies that process personal data [Webinar]. PennAITech. https://www.a2collective.ai/resources/ethical-responsibilities-for-companies-that-process-personal-data

Mello, M.M. & Cohen, I.G. (2025). Regulation of health and health care artificial intelligence. *JAMA, 333(*20): 1769-1770. https://doi.org/10.1001/jama.2025.3308

Mikhail, A. (2025, May 31). The reality of AI's promise to curb older adults' loneliness. *Fortune Well*. https://fortune.com/well/2025/05/31/ai-promise-older-adults-loneliness/.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. U.S. Department of Health and Human Services. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

New York State Office for the Aging (2024). Transforming care with AI for older adults: Two years of success of NY older adults living with ElliQ. [Report]. https://aging.ny.gov/system/files/documents/2024/10/transforming-care-for-older-adults-two-years-of-success-of-older-adults-with-elliq.pdf

Olympic Area Agency on Aging (2025). ElliQ pilot project report. https://www.o3a.org/files/2025/09/ELLIQ-PILOT-PROJECT-REPORT-2025.pdf

Parens, E. (2016). *Shaping our selves: On technology, flourishing, and a habit of thinking*. Oxford: Oxford University Press.

Perone, A.K., Abadir, P.M., Berlinger, N., Carey, J.R., Guest, M.A., Hass, A.J., Stephan, A.T., & Xie, B. (2026). Multidisciplinary perspectives on artificial intelligence in aging research

and education: evolving uses, ethics, and equity considerations in gerontology, *The Gerontologist* 66(4) https://doi.org/10.1093/geront/gnaf314

Phang, J., Lampe, M., Ahmad, L., Agarwal, S., Fang, C.M., Liu, A.R., Danry, V., Lee, E., Pataranutaporn, P., & Maes, P. (2025, March 21). Investigating affective use and emotional well-being on ChatGPT. *MIT Media Lab and Open AI report*. https://www.media.mit.edu/publications/investigating-affective-use-and-emotional-well-being-on-chatgpt/

Reuel, A. (2024). Chapter 3: Responsible AI. In *Artificial Intelligence Index Report 2024*. Stanford University Human-Centered AI. https://hai.stanford.edu/assets/files/hai_ai-index-report-2024_chapter3.pdf.

Saslow, E. (2026). To stay in her home, she let in an A.I. robot. *The New York Times.* February 12. https://www.nytimes.com/2026/02/12/us/elliq-ai-robot-senior-companion.html

Skuler, D. (2024). Why we created an AI code of ethics and why you should consider one for your company. Forbes Technology Council. *Forbes*. https://www.forbes.com/councils/forbestechcouncil/2024/10/29/why-we-created-an-ai-code-of-ethics-and-why-you-should-consider-one-for-your-company/

Symons X. & Savulescu, J. (2026). On old age: A relational account of agency and meaning in later life. *Hastings Center Report, 56*(1), 24-31. https://doi.org/10.1002/hast.5022

Taylor, C.M. (1991). *The ethics of authenticity.* Harvard University Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, Volume LIX(236): 433–460, https://doi.org/10.1093/mind/LIX.236.433

Turkle, S. (2017). *Alone together.* Revised and expanded third edition. Basic Books.

Vasani, S. (2026, May 17). ElliQ is a surprisingly helpful companion robot for older adults. *The Verge.* https://www.theverge.com/gadgets/928806/elliq-intuition-robotics-hands-on

Wallace, B. (2026, July 5). The Revenge of the Philosophy Majors. *The New York Times*. https://www.nytimes.com/2026/07/05/business/philosophy-majors-ai-jobs.html

Wexler, A. (2023). Ethical considerations for wearable devices and AI applications [Webinar]. PennAITech. https://youtu.be/KBBRQGGT_AY

World Health Organization Health Ethics & Governance Unit. (2021). Ethics and governance of artificial intelligence for health: WHO guidance. https://www.who.int/publications/i/item/9789240029200?utm