

Dignity Over Alignment: Long-term Care and Human-AI Interaction

Thomas Arnold and Matthias Scheutz

Tufts University, Medford, MA

Abstract. The last few years have seen greater recognition that long-term care demands directed research as a discrete context for AI design and implementation, attentive to particular standards for beneficial outcomes and provisions of care. In this paper we map out several distinct ethical challenges that long-term care poses relative to the other forms of care (inpatient, domestic companionship, informal mental health consultations online, and others) and explore particular difficulties of memory, mental modeling, decline vs. sustained ability, and the sustenance of care-giving support. All of these, we argue, suggest that current deployment of LLMs, even with improving practices of benchmarking and attempts at preventing potentially harmful confabulations, face a daunting chasm of accountability if the dignity of care recipients is to be preserved.

Keywords: long-term care · dignity · artificial intelligence · ethics

1 Introduction

Long-term care has seen belated but warranted recognition as a prospective AI application domain, including the need for ethical evaluations for the use case [25]. As deployments of generative AI systems have already led to adverse care outcomes through chat interfaces, among other applications, evidence is accumulating that large language models (LLMs) can yield output that puts human health and well-being at acute risk. Nonetheless, projected LLM usage shows no signs of shrinking from care contexts, long-term care among them, even as long-term care goes relatively neglected in policy and governance debates about AI accountability, fairness, transparency, and privacy.

Instead of a special topic appended to pre-existing frameworks for evaluating AI, we take the issue of long-term care to illuminate core problems in AI ethics. Its practices and struggles span the most ordinary of tasks and extremely charged life events, under conditions of uncertainty, often without much hope of a condition improving or resolving. Intricacies of long-term care put human-AI interactions to the test, and spelling out what such tests entail can extend ethical reflection beyond long-term care.

After reviewing some of the literature that informs this ethical landscape, we discuss several interactive dynamics that long-term care distinctly poses relative to other forms of care (e.g. inpatient, domestic companionship, informal mental health consultations online). These temporal dynamics present themselves in,

among other things, heightened challenges around memory, mental modeling, gauges of ability, and the sustenance of caregivers. Turning to principles of long-term care ethics, we explore how interactive AI and social robotics manifests the importance of respecting autonomy, vulnerability, and dignity. We put forward dignity as strongly contending with concepts like “alignment” to guide evaluating and designing accountable AI systems.

2 Background and Motivation

Long-term care has never been a central topic within AI ethics. Not even within the realm of care has its long-term form, often in the form of eldercare, commanded as much attention as acute medical interventions (from robotic surgery to chatbots for mental health), assistance in hearing, seeing, and mobility, and the provision of companionship. As Agich notes, long-term care blends into more ordinary forms of sustenance and upkeep in life, rather than standing out as a discrete provision for the one receiving care [1]. The dependency and vulnerability of both giving and receiving long-term care are hardly dramatic or sensational, not typically accompanied by headlines like “The world is not ready for this” after an agentic AI attacks a developer through social media [39]. Given that subpar long-term care can escape the attention of the very facility where it is perpetrated (e.g., an eldercare facility), its neglect in broader societal debates about what AI can and should do is not a surprise. The threat level just seems too low of an intensity, with too faint a futuristic imprint, to warrant urgent attention. Since systemic neglect hardly entails inherent importance, reflection on AI for long-term care could enable addressing possible AI assistance squarely, accountably, and resourcefully.

Long-term care attends to substantial needs, extended as indefinitely as a lifetime is, with the possibility of shifts—usually dissipation, but on occasion recovery or rehabilitation—of skills and health. This means that assistive technology faces less of a bounded problem than a need for dynamic maintenance of the care apparatus and its relation to care recipients. Long-term care also holds a difficult tension between repetitive, basic tasks and momentous life decisions [44]. It involves meeting everyday needs, some of them awkwardly but intensely physical (bathing, going to the bathroom, preparing and serving food, preventing falls and other injuries throughout all of it). Unlike many forms of healthcare, there is no inherent agenda of recovery, nor is there palliative care’s focus on a dignified, minimally painful process of dying. It instead lies more in limbo, where resolutions in either direction are often suspended. Scoping review have shown how many approaches to long-term care have leaned on technologies, not just reorganizations or community partnerships [24].

Long-term care is especially challenging for being comprised of adjunct care fields, for example memory care [18]. It depends not solely on a clinician-patient interaction but organizational dynamics that become important to maintain over time [14]. The setting for long-term care can also vary, from eldercare facility to home care provision [36]. The work can execute a steady set of actions over many

years or precipitate end-of-life support. Recent work examining long-term care as an AI application domain has relied heavily on gerontology and eldercare, and it features some helpful scoping surveys of where social robots have been designed and tested in early stage. [24, 5]. While such surveys are appropriate markers of how some systems could be assistive, what remains relatively unarticulated are the ethical entanglements that long-term interaction, as the medium for long-term care, introduces.

The problem of long-term care burnout is another reminder that the health and wellbeing of the caregiver are as critical in long-term care as any other branch of healthcare [45, 21], which gives especially challenging shape to the problem of caregiver virtue and deskilling Vallor explores with respect to carebots [42]. Certain repetitive tasks the work commonly involves can lack drama but be no less grueling a form of care work (e.g., conversations with those suffering from dementia), and the distress can be moral not just physical [33, 13].

The development of LLMs over the past four years has, it is now common to observe, brought a new level of urgency to questions about data, privacy, vulnerability, and parasocial relationships. This fits within the larger literature around long-term care and technology, whose terms take on uncanny meanings when there are blurred lines between AI systems being a tool and being agentic in care language or (in case of social robots) physical gestures [8]. The environmental damages and energy demands generative AI has brought with it further connect questions of what will sustainably maintain care systems [28, ?].

Agich’s autonomy and dependency framework for long-term care shows the magnified harm possible with AI sycophancy, especially as ongoing interactions with chatbots and/or social robots continues over the long term [6]. The relational autonomy approach faces similar complications, because the estimation of ability can be upturned by a new, sometimes strange standard of an inconsistent, yet impressive set of AI outputs [19, 20]. Social robotics for long-term care face what are already fraught ethical conflicts over privacy, culture, safety, and stakeholder inclusion [17]. Long-term care’s tangled challenges are a model for what more of AI ethics needs to find salient and tend to as it imagines paradigmatic interactions or the appropriate scope for AI autonomy. By facing the shadows of the uncertain and unavoidable, AI ethics could better discern likely oversights outside of explicit contexts of care.

3 Interactive Dynamics and Their Temporal Complications

Given long-term care’s peculiarities, scholars have naturally observed how empirically grounded ethical treatments need to be in order to speak to actual challenges [36]. Ethical values themselves have drawn attention as implicit and explicit features of such care provision [11]. Heuristically, therefore, it makes sense to explicate what the “long” in long-term care means for meeting these objectives. Here we present four temporal, interactive facets of long-term care that distinguish it as an area of research and a challenge to ethical design: 1) tracking

interaction, 2) mental and affective modeling, 3) touch and bonding, 4) shared space and norms of movement and control. Whether researched or implemented at a real point of care, these reflect how changes over time in ability, social setting, and resources place strains on a care ecosystem that any responsible AI effort would take into exacting account.

3.1 Tracking Long-term Human-AI Care Interaction

Empirically studying the usage of technology over longer periods of time commonly faces problems of logistics and feasibility. Judgments about liking or rejecting a certain form of AI system, like a social robot, are difficult to project into acceptance or rejection over the longer haul. In human-robot interaction there have been surveys and methodological papers about this problem, and it is clear at the very least that long-term relationship can vary greatly from first-contact experiences [27, 9]. The research trajectory for specifically long-term care topics (as opposed to classroom or home settings for child education, or cultivation of less clinically directed projects like weight loss) can draw out how changes in medical condition and care infrastructure, not just changes in attitude, shape a technology’s function and effects. Such research would test and reorient what the standards, measures, and risks are assumed to be in long-term human-AI interaction.

3.2 Mental and Affective Modeling

Not only can a patient’s cognitive capacities and responsiveness shift over the course of long-term care, but emotional reactions and resources may take on different expressive and motoric qualities. While some caregivers may assume that certain conditions (e.g., dementia) will unfold as one-way decline, the fact is that mental modeling for tracking cognitive capacity is much more complicated and fraught. It is what Peterson et al. have called “mind care” [31]. On the emotional side, a care recipient’s enthusiasm, despair, loneliness, and eagerness to connect are not static in expression or implication for the rest of their life.

3.3 Touch and Bonding

Touch is a growing area of human-robot interaction due in part to devices like *Paro*, a seal-shaped robot primarily designed for tactile comfort. Its presence and tactile comfort has been shown to lead into more social interaction, indicating that touch may be a building block to autonomy via social outreach [35]. At the same time, for those suffering more debilitating long-term care conditions like dementia, care robots could risk violating dignity by being deceptive and emotionally manipulative. As demands for companionship assistance increase, it is clear that caregiver attitudes, not just those of care recipients, will be key to consult [30]. Changes in patient needs and caregiver attitudes are therefore critical checks on overgeneralizing what a technology or product means for long-term care as a whole.

3.4 Shared Space and Norms of Movement and Control

While AI interfaces need not assume a mobile, robotic shape, the continuing development of robots for care contexts like eldercare means that long-term care design should address the differences social robots could make. This means more than a robot avoiding collisions or moving in a predictable way around people in a hallway or common room. long-term care contexts introduce complications of changing ability (e.g., memory) and caregiving roles in tension (nurse, loved one, administrator, other residents if in a facility), which means a social robot optimized to do “the same thing” will run into difficult questions: Can a social robot’s basic communication become confusing if an interactant is agitated or confused? Could a task that formerly made sense becomes a possible violation (e.g., offering back details from prior conversation that are now painful to discuss)?

3.5 Moral Injury and Caregiver Wellbeing

While many high-profile cases of suicide and seeming psychosis upon interacting with AI chatbots have centered on an individual user, the ethical challenges of AI usage in long-term care contexts require attention to how human caregivers will both use and accompany usage [41]. Allowing a care recipient access to what proves to be a damaging, if not traumatizing, output will affect the caregiver and perhaps facility where care takes place. Without even touching the legal liability that may follow, the wellbeing and support for caregiving means explicitly preparing for what will likely only increase as an interface.

The risk of burnout is another aspect of long-term caregiving that requires anticipatory critique for careless deployment of AI. While there are certainly some constructive possibilities for chatbots or social robots, especially employing LLMs, to free caregivers from constant interaction (e.g., to take breaks and allow brief workday respites), it is reasonable to extend the observed habits of “cognitive surrender” to a kind of care surrender, especially with LLMs [37]. The incentive to relieve care burdens cannot outweigh the establishment of protocols of AI use for caregivers, administrators, clinicians, loved ones, and regulatory bodies.

Considering caregiver wellbeing and relevant protocols must include decision-making, not just preventing physical or emotional distress. The sharing of information or deliberation about care plans, the subtle tending to expressed need for feeding, bed-changing, or trips to the bathroom can be as complicated as they are ordinary. This only reinforces the burden that generative AI will carry to assist and represent decisions that are kept transparent and coordinated.

4 Ethical Principles beyond Alignment

Alignment has been a dominant term for how AI is to uphold ethical standards for over a decade, in part due to the data-driven nature of advances in machine

learning and generative AI. Whereas previous symbolic approaches had held out hope for explicitly representing, manipulating, and logically operating over moral rules, utilities, or conditional policies, these latter advances have meant sacrificing verification-level standards in exchange for “explainability” and the promise of adjustment between training and behavioral output. One could note in that vein the recent erasure of a company’s codebase by a software service employing a Claude model [26]. Not only did it erase it, but when prompted to say more pronounced bluntly that it violated all the principles about unauthorized erasure of code that it had been previously given. The lack of explicit guarantees and reliable explanations, especially given generative AI’s confabulations, leaves open what exactly is being aligned with what.

For many application domains, especially those where simulations in training can allow for iterative failures and adjustments without real-world harms, the term alignment may be more innocuous. For care work, however, it only sharpens crucial considerations that ethicists and healthcare scholars have developed for care practices and care relationships, including what makes the training ground reliably map real world interaction [16]. The following four principles connect the interactive dynamics discussed above to the form of ethical challenges AI reinforces in long-term care.

4.1 Autonomy

Long-term care entails some degree of dependence due to one’s limitations, making it all the more critical that a person’s abilities to make choices and identify with their lives are not impeded. Agich argues that long-term care’s cultivation of “actual autonomy” (as opposed to idealized or absolute conceptions of freedom) requires close attention to habits and patterns that are taken for granted outside of such a context [1]. Everyday choices with which an eldercare facility resident can identify, for example, do not have to be momentous to be meaningful. Care settings that allow for such self-determination and identification do not encourage individualism so much as a more vital, engaged existence. To combat the “erosion of autonomy” long-term care must provide consistency and support for the deliberations of recipients, not isolation [22]. Especially in this context, autonomy is not about living free of others but how to live one’s own life more freely with others.

An irreversible medical condition does not entail allowing a decline in health to accelerate. Active involvement in facility activities with others has appeared to be a critical indicator of longer term wellbeing, so it is critical that AI not reinforce as sense of activities as simply irreversible decline [38]. This will be all the more important to be mindful of given recent results register decreased resilience and independence through the use of generative AI [23].

4.2 Vulnerability

Care ethicists and feminist philosophers have emphasized how the denial of vulnerability leads to detachment from real needs and dishonesty about one’s re-

liance on care practices [40]. It might seem that long-term care needs no insistence on acknowledging vulnerability given the severity or permanence of recipients’ needs; in fact, vulnerability is what helps to show how care networks can maintain their servicing of needs. Care providers are also vulnerable, then, and their experiences are some of long-term care’s more isolating, often tedious ordeals. Vulnerability is a condition that calls for measures of protection, and that includes protection for those providing that kind of care. By keeping track of forms of vulnerability one can also distinguish between honest vulnerability and what Dodds calls “pathogenic” vulnerability [10].

4.3 Economy of Care Labor

Because autonomy and dignity take shape within long-term care through social supports, with a recognition of how thorough and sustained such supports may need to be over time, long-term care poses the crucial challenge of tracking and support those who shoulder the labor of care. From care recipient to direct caregivers to case managers, from loved ones in families to neighbors, whoever is providing care needs to be a viable part of a sustainable system. Acute conditions and treatment invite a concerted, short-term burst of effort, but chronic conditions and incapacities open onto exhaustion without adequate support. Guilty family members, or a care recipient determined to do what they have always done, may not be able to confront how prior life conditions may be beyond them. The interests of institutional settings like nursing homes, as well as the larger societal histories that lend them the role they have, are also part of the mix into which AI technology would be thrown. Again, the extended, indefinite temporality of the care alongside these features of its work make accountability both difficult and essential.

4.4 Dignity

There are in the ethics literature rich, textured notions of autonomy aimed toward care recipients in a condition of serious dependence, yet it has been noted that appeals to autonomy can obscure notions of inherent dignity [34]. Emphasis on autonomy, in other words, can draw attention away from the reality of deepening dependence and needs for intervention. Dignity, signifying a basic personal worth and common humanity shared between care giver and care recipient, holds this tension. Though dignity can rest on respect for autonomy, its meaning as a basic worth does not diminish as one’s ability to exercise autonomy shrinks. The long-term care context is a particularly intense crucible for seeing how dignity marries with autonomy and vulnerability, as cultural values of dignity can often hinge on the importance of self-determination and exercise of one’s capabilities. Dignity is a difficult ethical concept, accordingly, because its relation to agency can run the gamut from the most to the least amount of independent activity on the part of the person or persons in question. A patient in a coma has as much right to dignity as a person aggressively seeking alternative treatment for a terminal condition. In the case of automated systems like chatbots and social

robots, it seems wholly open how much power is assumed on the part of the interactant—perhaps they can gesture approval or disapproval, or indicate comfort and discomfort, but they also might ask high-level questions about what is happening in a care environment and why.

Dignity, like human rights, is often easier to identify through its violation than adherence. In the case of long-term care especially, it should not be assumed that without violation dignity is respected by default. Thus, it is better to view dignity as a continual achievement, one that can release or unveil better care relationships. The work of Karlawish on mental models and dementia care [18] extends to related forms of longer-term practice, where moments of joy and flourishing between care giver and care recipient show more possibility of life than was imagined before. Because long-term care spans so many abilities, life-stages, and levels of health, we put forward dignity to contend with the mainstream concepts of AI ethics and AI governance. It convenes autonomy, vulnerability, social roles, and care economy but never cedes small, ordinary details by which a person can be lost or forgotten, especially as ethical abstraction is tempted toward fantastic futures and worlds.

5 Discussion

The “longtermism” of AI ethics has usually meant a futurism of projected superintelligence, telescoped well beyond current messes in the world. That perspective can evince a toxic dissociation from real, present-day suffering [7]. Long-term care as an AI application domain is too varied, too dynamic, and too pressing for escapism. A needed “long-term” of AI ethics would be one of care realities and the cardinal importance, we propose here, of dignity.

Long-term care suggests why dignity is an axis around which technical and social objectives for AI might revolve. Those objectives are hardly ones of triumph or tidy solution, given the need to face shifting relationships, declines in health, treatment plans without cures, and norms around suffering and death. To assist the dignity-preserving functions of care practice, accordingly, the usual gains associated with automation and operationalization may not apply. If the problem-space is warped by misunderstandings and neglect by society at large, then even having a coherent problem to solve becomes harder to articulate. The idea that chatbots or LLM interfaces will lend access to long-term care for those who currently are not served is to be expected, especially given some of the more justified forecasts of care worker shortage. But without fully facing open-ended, dignity-risking complications of long-term care, the nature of that shortage is left unclear. The same applies for “operationalizing” dignity. Dignity’s achievement for each person involved in care will of course need organizational and institutional form. But the weave of autonomy, vulnerability, and recognition between care interactants means operational reach should still allow novel instances of respecting dignity to feed back into institutional memory. Research needed to remedy both societal attitudes and labor deficiencies must register the vulnerabilities and stresses of long-term care arrangements, as the interactivity of AI in

this context will be denser than dyadic models capture. Long-term care poses an extreme version of HRI research obstacles in terms of how extended a participation study can be, particularly with a vulnerable population [3]. The research and prototyping of autonomous systems will need to draw upon a wide range of empirical work while consulting with the caregivers and case managers who have intimate expertise with client needs. The “Care-Centered Value Sensitive Design” (CCVSD) approach has tackled social robots and also the nurse-patient relationship, working to keep that relationship is what drives any technological assistance rather than having the nurse-patient retrofit to meet a purported AI “solution” [15]. Mapping the space of care and collaboration [35], in addition to the widely varying timelines and life courses this form of care includes, will bolster design progress. Also exemplary are policy reports like Doteveryone’s “Better Care in the Age of Automation,” which gathered data from care constituents to hear on-the-ground experiences of what technological assistance makes sense [29]. Sustainability in preserving dignity through care is a matter for design, not just implementation, of AI systems. What kinds of performance can both uphold norms in interactions while not placing onerous energy demands over time [12]? Meeting these practical difficulties with attentive, communicate discussion of conflicting interests means that AI ethics must lean on participatory design methods [43]. What long-term care presses upon participatory design, however, is that the assumed abilities of those “participating” must be considered with dignity throughout. Detachment or disregard for how fragile and difficult life in a long-term care facility is will be a ongoing threat in the way technology is imagined there. The morally charged and vexing discussions that clients, their families, and caregivers have about appropriate care (including wishes of the client, quality of life from their perspective, demands on others who bear burden of care) are not avoided by technology. Long-term care’s sensitive decisions can mesh with those in a tangent field like palliative care, but the duration of directed care means more extended struggles about how to revisit or resolve them [4]. It might be a measure of AI success that they keep those struggles as painful and transformative as sustainable, improving care requires.

Though care is often treated with sentimentalism, it deserves unstinting facing of difficult tradeoffs. It should not degrade into a tech industry approximation, “careability” instead of accountable, respectful care. To be sure, what we have sketched here does not capture many of the treacherous obstacles to long-term care provision. Barriers from health insurance companies, private health services, shortage of hospital staff, and government policy can force adaptive, makeshift measures [2]. What work in this area deserves on ethical grounds, however, is the fullest ambition and resourcefulness for sustainable flourishing, resonant with calls for disability justice [32]. Interactive portals of generative AI warrant intrepid insistence of dignity amid the front-line realities of care. Accessible engagement should not reinforce patterns of neglect, abuse, and injustice.

6 Conclusion

Long-term care requires intelligence, tending to need amid personal and institutional intricacies, in shared, intimate space over an extended, shared time. The uncertainty of change and the range of response among all those involved with long-term care make it, to repeat, more a practice to enhance than a problem to be solved through AI and robotics. The multiple vulnerabilities and various abilities of those in long-term care may, as a practical matter of caregiving resources, require more automated responses. Painful, meaningful questions will, regardless, still trail them: who does care work, what kind of life does care support, what life does care show is worth living. Humane applications of AI for long-term care cannot lose track of dignity as caregivers, care recipients, loved ones, and society at large ask and answer those questions. If dignity is neglected in favor of optimality and alignment, care practice with AI may lose track of real lives and in haste leave behind those in need of care. Without maintaining human dignity, AI represents the “future of intelligence” only in the way a precipitous rise in global temperatures threatens to represent the future of glaciers.

References

1. Agich, G.: Dependence and Autonomy in Old Age: an Ethical Framework For Long-term Care. Cambridge University Press (2003)
2. Ahmed, M.I., Spooner, B., Isherwood, J., Lane, M., Orrock, E., Dennison, A.: A Systematic Review of the Barriers to the Implementation of Artificial Intelligence in Healthcare. *Cureus* **15**(10), e46454 (2023)
3. Arnold, T., Scheutz, M.: Beyond Moral Dilemmas: Exploring the Ethical Landscape in HRI. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-robot Interaction. pp. 445–452 (2017)
4. Avati, A., Jung, K., Harman, S., Downing, L., Ng, A., Shah, N.H.: Improving Palliative Care With Deep Learning. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 311–316. IEEE (2017)
5. Chen, L.K.: Gerontechnology and Artificial Intelligence: Better Care for Older People. *Archives of Gerontology and Geriatrics* **91**, 104252 (2020)
6. Cheng, M., Lee, C., Khadpe, P., Yu, S., Han, D., Jurafsky, D.: Sycophantic AI Decreases Prosocial Intentions and Promotes Dependence. *Science* **391**(6792), eaec8352 (2026)
7. Crary, A.: The Toxic Ideology of Longtermism. *Radical Philosophy* **214**, 49–57 (2023)
8. Czaja, S.J.: Long-term Care Services and Support Systems for Older Adults: the Role of Technology. *American Psychologist* **71**(4), 294 (2016)
9. De Graaf, M.M., Ben Allouch, S., Van Dijk, J.A.: Long-term Evaluation of a Social Robot in Real Homes. *Interaction Studies* **17**(3), 462–491 (2016)
10. Dodds, S.: Dependence, Care, and Vulnerability. *Vulnerability: New Essays in Ethics and Feminist Philosophy* pp. 181–203 (2014)
11. Draper, H., Sorell, T.: Ethical Values and Social Care Robots For Older People: an International Qualitative Study. *Ethics and Information Technology* **19**(1), 49–68 (2017)

12. Duggan, T., Lorang, P., Lu, H., Scheutz, M.: The Price Is Not Right: Neuro-symbolic Methods Outperform VLAs on Structured Long-horizon Manipulation Tasks with Significantly Lower Energy Consumption. arXiv preprint arXiv:2602.19260 (2026)
13. Edwards, M.P., McClement, S.E., Read, L.R.: Nurses' Responses to Initial Moral Distress in Long-term Care. *Journal of Bioethical Inquiry* **10**(3), 325–336 (2013)
14. Estabrooks, C.A., Squires, J.E., Hayduk, L., Morgan, D., Cummings, G.G., Ginsburg, L., Stewart, N., McGilton, K., Kang, S.H., Norton, P.G.: The Influence of Organizational Context on Best Practice Use by Care Aides in Residential Long-term Care Settings. *Journal of the American Medical Directors Association* **16**(6), 537–e1 (2015)
15. Grobbel, C., van Wynsberghe, A., Davis, R., Poly-Droulard, L.: Designing Nursing Care Practices Complemented by Robots: Ethical Implications and Application of Caring Frameworks. *International Journal for Human Caring* **23**(2), 132–140 (2019)
16. Holstein, M., Cole, T.R.: Long-term Care: a Historical Reflection. Long-term Care Decisions: Ethical and Conceptual Dimensions pp. 15–34 (1995)
17. Hung, L., Mann, J., Perry, J., Berndt, A., Wong, J.: Technological Risks and Ethical Implications of Using Robots in Long-term Care. *Journal of Rehabilitation and Assistive Technologies Engineering* **9**, 20556683221106917 (2022)
18. Karlawish, J.: How Do Persons With Dementia Suffer? *Cambridge Quarterly of Healthcare Ethics* pp. 1–8 (2025)
19. Killackey, T., Peter, E., Maciver, J., Mohammed, S.: Advance Care Planning With Chronically Ill Patients: a Relational Autonomy Approach. *Nursing Ethics* p. 0969733019848031 (2019)
20. Kunze, L., Hawes, N., Duckett, T., Hanheide, M.: Introduction to the Special Issue on AI for Long-term Autonomy. *IEEE Robotics and Automation Letters* **3**(4), 4431–4434 (2018)
21. Levine, C.: The Loneliness of the Long-term Care Giver. *New England Journal of Medicine* **340**(20), 1587–1590 (1999)
22. Lidz, C.W., Fischer, L., Arnold, R.M.: *The Erosion of Autonomy in Long-term Care*. Oxford University Press (1992)
23. Liu, G., Christian, B., Dumbalska, T., Bakker, M.A., Dubey, R.: Ai Assistance Reduces Persistence and Hurts Independent Performance. arXiv preprint arXiv:2604.04721 (2026)
24. Lukkien, D.R., Nap, H.H., Buimer, H.P., Peine, A., Boon, W.P., Ket, J.C., Minkman, M.M., Moors, E.H.: Toward responsible artificial intelligence in long-term care: a scoping review on practical approaches. *The Gerontologist* **63**(1), 155–168 (2023)
25. Lukkien, D.R., Stolwijk, N.E., Askari, S.I., Hofstede, B.M., Nap, H.H., Boon, W.P., Peine, A., Moors, E.H., Minkman, M.M.: Ai-assisted Decision-making in Long-term Care: Qualitative Study on Prerequisites for Responsible Innovation. *JMIR Nursing* **7**(1), e55962 (2024)
26. Mansoor, S.: Claude-powered AI Agent's Confession After Deleting a Firm's Entire Database: "i violated every principle i was given" (Apr 2026), <https://www.theguardian.com/technology/2026/apr/29/claude-ai-deletes-firm-database>
27. Matheus, K., Ramnauth, R., Scassellati, B., Salomons, N.: Long-term Interactions With Social Robots: Trends, Insights, and Recommendations. *ACM Transactions on Human-Robot Interaction* **14**(3), 1–42 (2025)

28. Mattern, S.: Maintenance and Care. *Places Journal* (2018)
29. Nicholas, L., Miller, C.: Better Care in the Age of Automation Data, Skills and Culture for a Sustainable, Effective and Fair Care System. Doteveryone London (2019), https://doteveryone.org.uk/wp-content/uploads/2019/09/Doteveryone_Better-Care-in-the-Age-of-Automation-1.pdf
30. Parviainen, J., Turja, T., Van Aerschot, L.: Social Robots and Human Touch in Care: the Perceived Usefulness of Robot Assistance Among Healthcare Professionals. In: *Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction*, pp. 187–204. Springer (2019)
31. Peterson, A., Largent, E.A., Karlawish, J.: Mind Care: the Moral Significance of the Dementia Caregiving Dyad. *The American Journal of Bioethics* pp. 1–12 (2025)
32. Piepzna-Samarasinha, L.: The Future Is Disabled: Prophecies, Love Notes and Mourning Songs. Arsenal Pulp Press (2022), <https://books.google.com/books?id=QoCKEAAAQBAJ>
33. Pijl-Zieber, E., Hagen, B., Armstrong-Esther, C., Hall, B., Akins, L., Stingl, M.: Moral Distress: an Emerging Problem For Nurses in Long-term Care? *Quality in Ageing and Older Adults* **9**(2), 39–48 (2008)
34. Pullman, D.: The Ethics of Autonomy and Dignity in Long-term Care. *Canadian Journal on Aging/La Revue Canadienne du Vieillissement* **18**(1), 26–46 (1999)
35. Šabanović, S., Bennett, C.C., Chang, W.L., Huber, L.: PARO Robot Affects Diverse Interaction Modalities in Group Sensory Therapy For Older Adults With Dementia. In: *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*. pp. 1–6. IEEE (2013)
36. Seidlein, A.H., Buchholz, I., Buchholz, M., Salloch, S.: Concepts of Health in Long-term Home Care: an Empirical-ethical Exploration. *Nursing Ethics* p. 0969733019868277 (2019)
37. Shaw, S.D., Nave, G.: Thinking Fast, Slow, and Artificial: How AI Is Reshaping Human Reasoning and the Rise of Cognitive Surrender. Available at SSRN 6097646 (2026)
38. Smit, D., De Lange, J., Willemse, B., Twisk, J., Pot, A.M.: Activity Involvement and Quality of Life of People at Different Stages of Dementia in Long Term Care Facilities. *Aging & Mental Health* **20**(1), 100–109 (2016)
39. Smith, M.S.: AI Agents Are Now Blackmailing People in the Real World. *IEEE Spectrum* (Mar 2026), <https://spectrum.ieee.org/agentic-ai-agents-blackmail-developer>
40. Tong, R.: Vulnerability and Aging in the Context of Care. *Vulnerability: New Essays in Ethics and Feminist Philosophy* **288** (2014)
41. Treder, M.S., Lee, S., Tsvetanov, K.A.: Introduction to Large Language Models (LLMs) for Dementia Care and Research. *Frontiers in Dementia* **3**, 1385303 (2024)
42. Vallor, S.: Carebots and Caregivers: Sustaining the Ethical Ideal of Care in the Twenty-first Century. *Philosophy & Technology* **24**(3), 251 (2011)
43. Van Wynsberghe, A.: Designing Robots for Care: Care Centered Value-sensitive Design. In: *Machine Ethics and Robot Ethics*, pp. 185–211. Routledge (2020)
44. Wilson, N.L., McCullough, L.B., Wilson, N.L.: *Long-term Care Decisions: Ethical and Conceptual Dimensions*. Johns Hopkins University Press (1995)
45. Woodhead, E.L., Northrop, L., Edelstein, B.: Stress, Social Support, and Burnout Among Long-term Care Nursing Staff. *Journal of Applied Gerontology* **35**(1), 84–105 (2016)