

Application Fields and Innovative Technologies

Explainable and Trustworthy AI in Long-Term Care: A Socio-Technical Perspective M. Kampel, V. Gallistl, R. von Laufenberg. *Gerontechnology* 25(s)

Purpose Artificial intelligence (AI) is increasingly deployed in long-term care (LTC) to enhance safety, efficiency, and quality of care. Yet its introduction into settings marked by vulnerability, institutional constraint, and intimate bodily routines raises critical questions about explainability, trustworthiness, and the reconfiguration of care practices. This presentation examines three AI-based systems currently used in LTC: (1) ceiling-mounted depth sensors that continuously monitor residents' posture and movement and generate alerts when predefined fall patterns are detected; (2) social robots that identify residents via speech and facial recognition and initiate scripted verbal interactions; and (3) depth-camera-based scene-understanding systems installed in bathrooms that guide people with mild dementia step-by-step through toileting routines via real-time prompts [1]. While all three applications are considered, the analysis foregrounds automated toilet assistance as a focal case, using fall detection and social robotics as comparative examples to explore how AI reshapes autonomy, risk, and appropriate behaviour in institutional care. **Method** The study is based on a multi-perspective qualitative case study design in Austrian and international care facilities. Data collection combined in-depth interviews with residents, professional care workers, managers, relatives, and technology developers with extensive participant observation of everyday care routines and technology use. This allowed us to trace how AI systems are embedded in socio-technical assemblages that include institutional logics, regulatory constraints, architectural layouts, staffing levels, and material infrastructures, rather than operating as stand-alone tools [2]. The analysis focused on how different actors interpret, appropriate, or resist AI systems in practice, and how explainability, trust, and bias are enacted in concrete situations [3]. **Results and Discussion** Across all three cases, the analysis shows that AI systems subtly standardise what counts as risk, mobility, appropriate interaction, and "correct" performance of intimate bodily routines. Fall-detection sensors do not simply register events; they stabilise particular understandings of acceptable movement and legitimise decisions through the appearance of technological neutrality. Social robots reconfigure emotional labour and interaction by scripting encounters between residents and machines, while simultaneously demanding new forms of coordination and maintenance from staff. Automated toilet assistance systems promise to support autonomy and reduce workload by guiding people with mild dementia through toileting sequences, yet they also extend algorithmic observation into highly private spaces, impose normative scripts of bodily conduct, and redistribute responsibility between human carers and depth-camera-based scene understanding. Mis-detections and false prompts are experienced not merely as technical glitches but as matters of dignity, shame, and institutional risk management. The findings demonstrate that explainability in LTC cannot be reduced to technical transparency of algorithms. For care workers and residents, trust, reliability in practice, and the possibility to contest, circumvent, or creatively reinterpret AI outputs are more salient than access to model details. Users develop "folk theories" of how systems work, grounded in embodied and organisational experience. Bias likewise cannot be located solely in data or models; it emerges from the broader assemblage in which AI is deployed, including staffing regimes, institutional priorities, and built environments. The study concludes that AI systems in ageing and care must be designed and implemented with explicit attention to vulnerability, power relations, and the contested nature of autonomy in institutional settings. Rather than treating AI as a neutral optimisation tool, we argue that its role in reshaping the conditions under which dignity, relationality, and meaningful human interaction can be sustained in later life must be central to assessments of explainability and trustworthiness.

References

1. Ballester, I. Gall, M. Münzer, T. Kampel, M. Vision-Based Toilet Assistant for People with Dementia in Real-Life Situations. In *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)* (pp. 141-147). IEEE. [10.1109/PerComWorkshops59983.2024.10503526](https://doi.org/10.1109/PerComWorkshops59983.2024.10503526)
2. Gallistl, V., Bandy, M.U., Berridge, C., Grigorovich, A., Jarke, J., Mannheim, I., Marshall, B., Martin, W., Moreira, T., van Leersum, C.M., & Peine, A. (2024). Addressing the Black Box of AI—A Model and Research Agenda on the Co-constitution of Aging and Artificial Intelligence. *The Gerontologist*, 64.
3. Kontrus, V., von Laufenberg, R. Biased AI in Long-Term Care – It's more than just in the data. In *4th Annual Meeting of the Socio-Gerontechnology Network*, Vienna/Austria, 2024.

Keywords: Long-term care; Artificial intelligence; Explainability; Socio-technical systems; Dementia care

Affiliation: Technical University of Vienna, Computer Vision Lab; Favoritenstraße 9-11, A-1040 Vienna, Austria **Email:** martin.kampel@tuwien.ac.at

Acknowledgement The project is funded by the Erasmus+ Programme of the European Union (Grant Agreement No. 612655-EPP-1-2019-1-EL-EPPKA2-SSA)