ExpliCareNEXT: Human-Centered Explainable AI Solutions for Nursing Care

Julia Hermann[1,*], Sebastian van Ledden[2], Christian Meske[3], Enrico Löhrke[4], Edwin Naroska[5], Todor Dimitrov[6], Sabahat Ölcer[7,8] and Aysegül Dogangün[1]

[1]Institute Positive Computing, Ruhr West University of Applied Sciences, Lützowstraße 5, 46236 Bottrop, Germany
[3]Institute of Work Science, Ruhr University Bochum, Wasserstraße 221, 44799 Bochum, Germany
[4]inhaus GmbH, Grabenstraße 5, 47057 Duisburg, Germany
[5]Ambient Intelligence Laboratory, Niederrhein University of Applied Sciences, Reinarzstr. 49, 47805 Krefeld, Germany
[6]Anasoft Technology AG, Querenburger Str. 38, 44789 Bochum, Germany
[7]Institute of Computer Science, Ruhr West University of Applied Sciences, Lützowstraße 5, 46236 Bottrop, Germany
[8]Department of Palliative Medicine, University Medical Center Göttingen, University of Göttingen, Von-Siebold-Straße 3, 37075 Göttingen, Germany

Abstract

The increasing demand for nursing care services, coupled with a persistent shortage of skilled caregivers, underscores the urgent need for innovative technological solutions. ExpliCareNEXT addresses this challenge by integrating explainable artificial intelligence (XAI) into caregiving workflows to enhance efficiency, transparency, and trust. The project applies participatory design methods to develop adaptive, human-centered AI systems that support caregivers in their daily tasks while ensuring interpretability and ease of use. Initial findings from field studies in institutional care settings reveal barriers to technology adoption, including high documentation burdens, language barriers, and data privacy concerns. These insights inform the system design, ensuring that AI-driven recommendations align with caregivers' workflows and information needs. Our paper presents the project's objectives and methodological approach, emphasizing interdisciplinary collaboration as key strategies for bridging the gap between technical innovation and practical application. Challenges are explored, highlighting the importance of participatory approaches in fostering trust and usability in safety-critical domains.

Keywords

Participatory Design, Nursing Care, Caregivers, Explainable Artificial Intelligence, Human-Centered AI, Context-Sensitive Explanations

1. Introduction

The growing need for elderly care services, combined with a persistent shortage of qualified formal caregivers, poses substantial challenges for contemporary care systems. In Germany, approximately five million people currently require care, with 80% being supported at home by family members or outpatient care services [1, 2]. By 2035, this demand is expected to rise, leading to an estimated shortfall of 500,000 caregivers [3]. High workloads, frequent staff turnover, and widespread job dissatisfaction further exacerbate this issue, with 67% of caregivers considering leaving the profession [4]. These challenges underline the urgent need for innovative solutions that reduce the workload of care staff, allow them to focus on interpersonal tasks, and at the same time maintain high quality standards in care. Explainable Artificial Intelligence (XAI) offers promising opportunities to address these challenges by providing actionable, context-aware support. While existing IT-based solutions, such as electronic

Joint Proceedings of the ACM IUI Workshops 2025, March 24-27, 2025, Cagliari, Italy
*Corresponding author.
julia.hermann@hs-ruhrwest.de (J. Hermann); sebastian.vanledden@studium.fernuni-hagen.de (S. v. Ledden); christian.meske@ruhr-uni-bochum.de (C. Meske); loehrke@inhaus.de (E. Löhrke); edwin.naroska@hsnr.de (E. Naroska); dimitrov@technology.de (T. Dimitrov); sabahat.oelcer@hs-ruhrwest.de (S. Ölcer); ayseguel.doganguen@hs-ruhrwest.de (A. Dogangün)
0000-0003-0802-101X (J. Hermann); 0000-0003-3920-2841 (S. v. Ledden); 0000-0001-5637-9433 (C. Meske); 0000-0002-4287-3884 (E. Löhrke); 0009-0004-8742-4566 (E. Naroska); 0000-0002-3124-8039 (S. Ölcer); 0000-0002-4287-3884 (A. Dogangün)

care documentation, have demonstrated their potential in care environments, they often fail to achieve widespread adoption due to their complexity and lack of transparency [4, 5, 6].

In our project *ExpliCareNEXT* an initial requirement analysis in institutional care settings highlighted three major barriers to the adoption of systems based on artificial intelligence (AI): (1) excessive documentation requirements that increase the administrative burden on caregivers, (2) significant language barriers, as many care facilities rely on international staff, and (3) concerns regarding data privacy and security, particularly in the context of automated decision-making. These insights emphasize the necessity for AI-driven solutions that seamlessly integrate into caregiving workflows, offering multilingual support, clear decision rationales, and privacy-preserving mechanisms.

Our initial findings from first field studies indicate that caregivers are more likely to accept AI recommendations if they are presented in a comprehensible, situation-specific manner and are adaptable to individual caregiving styles. The acceptance in nursing care hinges on their seamless integration into caregivers' workflows, ensuring intuitive interfaces, and clear explanations of AI-driven recommendations [7, 8, 9, 10].

ExpliCareNEXT addresses these challenges by developing a context-adaptive, XAI-supported system tailored to nursing care settings. Its objectives are twofold: (1) to support supervisory staff in task prioritization and transparent decision-making processes, and (2) to assist caregivers with context-aware action recommendations that align with their skills, language preferences, and situational requirements. By leveraging data from smart living environments and electronic care documentation, the system aims to enhance efficiency, usability, and trust in caregiving workflows [11, 12].

Our paper presents the concepts, the methodological approach and the initial and expected future contributions of the project. It outlines the scientific and practical strategies used to meet the needs of diverse stakeholders and bridge the gap between technical feasibility and human-centered design.

2. Related Work

The "black-box" nature of many AI models has been identified as a significant barrier to adoption, particularly in high-stakes environments where errors can have severe consequences [7, 8]. Explainable Artificial Intelligence addresses the critical need for transparency in AI systems, particularly in domains where trust and accountability are paramount, such as elderly care. XAI methods aim to provide clear, interpretable explanations for AI decisions that enable not only developers but also domain experts or users to understand, trust and effectively manage these systems [6, 13].

Early research on explanation models identified key approaches for improving user understanding, including "trace or line of reasoning," "justification or support," and "terminological definitions" [14, 15]. These explanation types enable users to assess the validity and relevance of AI outputs, fostering trust and confidence in the system. More recent advancements, such as feature attribution methods (e.g., SHAP and LIME) and surrogate models, have further enhanced the interpretability of machine learning processes [16, 17]. However, balancing technical complexity with user accessibility remains a persistent challenge.

Trust in AI systems is closely linked to their ability to align with users' cognitive models and contextual needs. Studies show that personalized explanations tailored to users' expertise and situational demands significantly enhance perceived usefulness and acceptance [18, 19]. This is particularly relevant in nursing contexts, where caregivers operate under significant time constraints and administrative burdens. The effectiveness of XAI in such settings depends on its ability to provide explanations that are concise, actionable, and seamlessly integrated into existing workflows. Research highlights that caregivers often do not have the time to interpret lengthy, technical explanations, underscoring the need for context-aware and task-relevant XAI approaches [20, 21]. Research highlights the role of explainability not only in fostering trust but also in improving task efficiency and decision-making accuracy [5, 6]. Transparent systems that effectively communicate the rationale behind AI-generated recommendations empower caregivers to make informed decisions, ultimately leading to improved care outcomes [7, 8].

Advancements in natural language processing (NLP) and smart living technologies have significantly improved human-machine interaction, offering new opportunities to enhance care workflows. State-of-the-art speech recognition models such as Whisper enable precise multilingual conversion of speech into text and thus have the potential to support the diverse user groups in care environments [22]. Similarly, chatbots and AI-driven communication tools facilitate intuitive information retrieval and task management, though challenges such as hallucination and domain-specific inaccuracies persist [23, 24]. Efforts to integrate these technologies into comprehensive care systems have underscored the importance of transparency and data protection. Privacy-preserving methods, such as homomorphic encryption and federated learning, address critical concerns about data security while maintaining system functionality [12, 11]. These approaches are essential for fostering user trust and ensuring compliance with regulatory standards.

ExpliCareNEXT builds on this foundation by involving caregivers, technical experts, and other stakeholders in the design process, to address both technical and human-centered challenges. Based on the initial findings from our preliminary study [9, 10] and the requirements analysis carried out, we are concentrating on designing XAI explanations that are concise, contextual and directly applicable to the real decision-making processes of nursing staff. This approach ensures that the developed system is not only technically innovative but also practical and aligned with the needs of elderly care environments.

3. Methodology

The project applies Rapid Contextual Design [25], a participatory approach that integrates human-centered design principles. This method is tailored to develop a system specifically designed for caregivers in formal home and institutional care. By emphasizing inclusivity, transparency, and ethical considerations, the project unfolds across five interrelated phases, ensuring that the resulting AI system is not only technologically advanced but also aligned with the practical needs and ethical concerns of its end users, the domain experts.

The first phase of the project focused on capturing real-world requirements. Interviews with caregivers from institutional care settings provided detailed insights into documentation workflows, task prioritization, and daily challenges. Contextual inquiries [25] were conducted to capture contextual nuances that are often difficult to articulate in interviews and two initial workshops examined usability issues in AI-supported documentation and explored caregivers' expectations regarding transparency in automated recommendations. To further explore the mentioned documentation challenges, the existing care documentation systems were examined, with a focus on user-friendliness and data structuring [23, 26].

The project integrates multiple XAI methods to ensure transparency and interpretability. Following established approaches for practical XAI implementation [27], Local Interpretable Model-agnostic Explanations (LIME) [28] are used to provide localized explanations of AI-driven recommendations. Additionally, SHapley Additive exPlanations (SHAP) offer a quantitative breakdown of feature importance, helping caregivers understand the factors influencing automated decisions. Anchors [29] complement these methods by generating stable, rule-based explanations tailored to recurring care scenarios [30].

Explainability remains central to the system's development, with a focus on ensuring that AI-generated recommendations are actionable and comprehensible. The project applies established explanation approaches, such as "trace or line of reasoning," "justification or support," and "terminological definitions," to enhance transparency and usability [15, 14]. Adaptive learning mechanisms allow the AI to refine its recommendations based on user feedback, ensuring continuous improvement in decision support [11, 16].

At the same time, data privacy and security are key concerns, particularly given the sensitivity of caregiving data. The system architecture adheres to the General Data Protection Regulation (GDPR) and integrates privacy-preserving techniques, such as encryption and decentralized data processing, to ensure compliance with legal and ethical standards [31, 6]. A privacy-first design approach was adopted, enabling caregivers to retain control over their data while allowing AI-generated insights to

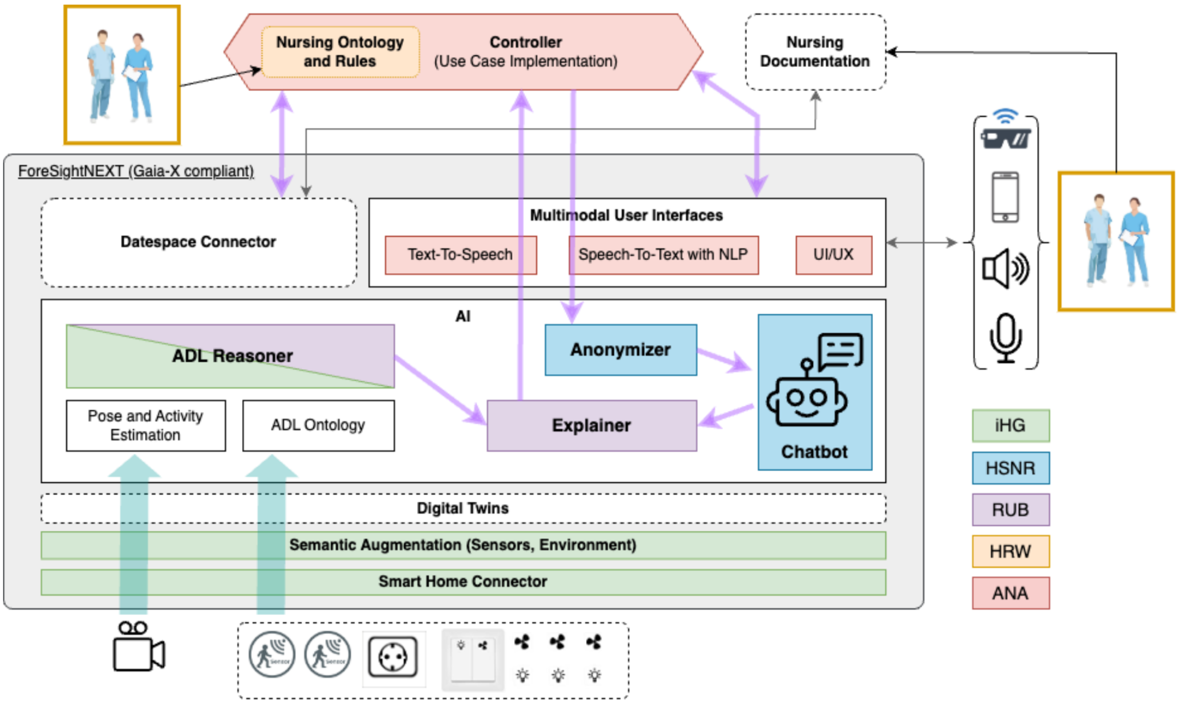

Figure 1: Architecture of ExpliCareNEXT and the roles of the project partners. The diagram illustrates the integration of AI, multimodal user interfaces, and digital twins to enhance nursing care. Partner contributions are denoted by color-coded labels: inhaus GmbH (iHG), Niederrhein University of Applied Sciences (HSNR), Ruhr University Bochum (RUB), Ruhr West University of Applied Sciences (HRW), and Anasoft Technology AG (ANA).

support but not replace human decision-making.

The evaluation phase combines controlled lab-based testing with field studies in real-world nursing care environments. Lab-based evaluations enable early usability testing, allowing researchers to identify interface issues and refine AI explanations [11]. These evaluations will be complemented by field studies that assess the system's impact on workflow efficiency, caregiver satisfaction, and decision-making support in practical settings [2, 4].

The iterative nature of our evaluation approach ensures that system refinements are continuously informed by feedback from domain experts and users. The findings from the requirements analysis play a crucial role in defining evaluation criteria, particularly regarding the acceptability and interpretability of AI-generated recommendations. Further iterations of co-creation workshops will continue to refine system usability, ultimately contributing to a model of explainable AI that is both practical and trustworthy in safety-critical caregiving contexts.

4. First Insights and Future Directions

Findings from the initial requirements analysis and co-creation workshops with caregivers confirm the necessity of AI-driven support in nursing care while also highlighting key challenges that must be addressed to ensure meaningful integration into daily practice.

Interviews and observational studies have revealed that caregivers face persistent difficulties related to documentation workload, language barriers, and skepticism toward automated decision-making. The initial results of the requirements analysis confirm our findings from our preliminary study [9, 10]: Caregivers experience significant burdens related to documentation requirements, which, instead of streamlining their work, often increase administrative complexity. Many existing digital solutions lack adequate usability, leading to frustration and inefficiencies. Furthermore, language barriers emerged as a critical issue, particularly among caregivers working in temporary employment or those recruited internationally. The inconsistent availability of multilingual interfaces contributes to misinterpretations, increasing the risk of errors in care provision. Additionally, skepticism toward AI-supported decision-making was observed, with caregivers expressing concerns about the reliability and interpretability of automated recommendations. Participants highlighted the need for context-sensitive explanations that align with their workflows, as well as mechanisms to verify and override AI-generated suggestions when necessary.

Building on these insights, the project has refined its approach to integrating explainability into caregiving workflows. LIME [28] and SHAP have been identified as promising techniques for generating localized and interpretable explanations, yet initial user feedback suggests that statistical explanations alone may not be sufficient. To enhance usability, the system incorporates contextualized rule-based explanations through Anchors [29], which enable caregivers to understand recurring decision patterns in a more intuitive way [30]. Furthermore, co-creation workshops have emphasized the importance of aligning AI-generated recommendations with caregivers' mental models and daily routines. As a result, system adaptations now include structured task prioritization, a user feedback mechanism for refining AI-generated suggestions, and a prototype for multimodal explanations that combine textual and visual elements to improve accessibility across diverse caregiver groups.

The project's ongoing technical advancements aim to bridge the gap between AI transparency and practical applicability. Multilingual interfaces have been prioritized to address language barriers identified during the initial analysis, ensuring that caregivers can interact with the system in their preferred language without compromising understanding or usability. Additionally, privacy-preserving mechanisms are continuously being refined to comply with ethical standards and data protection regulations, safeguarding sensitive user data while maintaining system functionality [31].

As the project progresses, upcoming evaluations will focus on validating these adaptations through further iterative testing. Field studies will examine how the system integrates into real-world caregiving environments, assessing its impact on workflow efficiency, decision-making processes, and overall caregiver satisfaction [2, 4]. Furthermore, research efforts will explore the scalability of the system to different caregiving contexts. In parallel, ethical considerations remain a central concern, with ongoing refinements in privacy protection and bias mitigation to ensure equitable AI-driven decision support.

5. Discussion

By aligning technical innovation with the practical needs of caregivers, *ExpliCareNEXT* contributes to the evolving discourse on human-centered AI, particularly in the context of trust and transparency [6, 7]. The integration of XAI into nursing care aims to support decision-making while maintaining transparency and usability. However, initial feedback from domain experts suggests that the way explanations are presented plays a crucial role in whether AI-generated recommendations are trusted or dismissed. While existing research emphasizes the importance of explainability for AI adoption [6, 7], our findings indicate that standard techniques such as feature attribution alone do not meet the needs of caregivers, who require more contextualized, rule-based explanations. This supports previous work showing that the success of XAI depends on aligning system-generated outputs with user expectations and domain knowledge [32].

Despite its theoretical advantages, explainability alone does not guarantee AI acceptance in safety-critical domains such as care. Trust in AI recommendations remains contingent on additional factors, including perceived accuracy, adaptability to individual workflows, and seamless integration into existing processes.

The use of participatory methods in AI system design has been advocated, yet their application in XAI development for caregiving remains limited. Our experience highlights a key challenge: the translation of participatory insights into technically viable, scalable AI explanations. While co-creation workshops allow caregivers to articulate their needs, defining a structured pathway to integrate these insights into AI model development requires additional methodological considerations.

A recurring issue in our process is the divergence between caregivers' conceptual models and the way AI systems generate explanations. While caregivers expect deterministic, scenario-based reasoning, AI explanations are often probabilistic, requiring additional interpretation. Similar findings have been reported in research on human-AI interaction, where the effectiveness of XAI is contingent on matching explanation formats to users' cognitive models [33]. This raises the question of whether domain-specific explanation strategies need to be formalized to bridge this conceptual gap.

Language accessibility has emerged as a more significant barrier than initially anticipated. Our findings suggest that language barriers extend beyond system interaction. Caregivers who are not fluent in the system's primary language may struggle to critically assess AI-generated recommendations, potentially leading to blind reliance or rejection of the technology. This raises important ethical considerations. If explainability is not tailored to different backgrounds, it risks reinforcing existing power asymmetries in care, as workers with limited technical or language skills may have less influence on decision-making. Future research needs to explore adaptive explanatory models to ensure that transparency mechanisms are effective for diverse user groups.

The insights gained from *ExpliCareNEXT* contribute to the broader discourse on XAI in safety-critical applications. Our findings align with existing work emphasizing that explainability must go beyond algorithmic transparency and instead be designed to enhance users' ability to act upon AI-generated recommendations [34]. A major open question remains: How can AI-generated explanations be evaluated for effectiveness beyond subjective user trust? While many studies focus on perceived trustworthiness, fewer explore the impact of explainability on actual decision-making accuracy and efficiency. This presents an opportunity for future work to develop evaluation metrics that link XAI effectiveness to measurable improvements in professional workflows.

Future work also needs to address the long-term impact of AI explanations on caregiver behavior and decision-making. In addition, scalability remains a critical issue: the methods and findings developed in this project need to be adapted to larger, more diverse care situations in order to assess their generalizability. The principles and methods developed in *ExpliCareNEXT* could also be adapted for cross-domain applications in other safety-critical areas, such as medical diagnostics or industrial automation.

6. Conclusion and Outlook

The *ExpliCareNEXT* project presents an approach to leveraging explainable AI (XAI) to address the increasing demands of formal home-based and institutional nursing care. By integrating participatory design methods and advanced AI technologies, the project highlights the importance of human-centered innovation in safety-critical domains. However, findings also emphasize challenges in translating user expectations into technically viable and scalable XAI solutions.

A key contribution of *ExpliCareNEXT* lies in the integration of context-aware explainability, ensuring that AI-generated recommendations remain actionable, transparent, and adaptable to the dynamic needs of caregiving tasks. The human-centered approach, grounded in Rapid Contextual Design [25], demonstrates how participatory methods such as contextual inquiry and work modeling can align technical development with real-world user needs. Initial results have highlighted pain points in current care workflows, particularly in relation to the usability of existing digital tools and the need for clear, language- and context-adaptive AI explanations.

In addition, the project focuses on ethical AI practices, including privacy-preserving data handling, informed consent and inclusive design. By ensuring that AI is used responsibly and effectively, this approach can improve the care of older people and complement the human aspects of care. These papers not only address the domain specific challenges, but also offer transferable insights for the development of AI systems in other challenging environments.

In the next project step, evaluation under real-life conditions will be crucial to validate the system's impact on caregivers' workflows, satisfaction and overall quality of care through field deployment and impact assessment. The next phases will also focus on refining explanatory strategies, especially in the context of AI-supported decision-making in care environments. Initial results indicate that nurses prefer explanations that provide rule-based, scenario-specific reasoning rather than abstract attributions of features.

The insights gained from *ExpliCareNEXT* highlight the importance of human-centered XAI in promoting trust, usability and impact in the real world. Our results show that the acceptance of AI in elderly care is closely related to the ability of caregivers to understand, interpret and modify the explanations and recommendations generated by the system. This underscores the need for explanatory mechanisms that prioritize actionable insights over purely algorithmic transparency.

As the field of XAI continues to evolve, *ExpliCareNEXT* serves as a case study for bridging the gap between technical potential and societal needs. The approach of *ExpliCareNEXT* is an example of how interdisciplinary collaboration and participatory methods can drive the development of AI systems that are not only technologically advanced, but also very close to the needs of the people they serve.

Acknowledgments

This research was funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) as part of the project ExpliCareNEXT (FKZ 01MK24003B).

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT-4 in order to: Grammar and spelling check. After using these tool, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] S. Bundesamt, Pflege im Fokus: Statistiken zu Pflegebedürftigen in Deutschland, Technical Report, Statistisches Bundesamt, 2022.

[2] A. Tutmann, Die Zukunft der häuslichen Pflege: Herausforderungen und Perspektiven, Technical Report, 2023.

[3] H. Rothgang, Pflege 2035: Prognosen für den fachkräftemangel, Gesundheit und Gesellschaft 11 (2022) 34–49.

[4] G. Braeseke, C. Pflug, T. Tisch, L. Wentz, U. Pörschmann-Schreiber, H. Kulas, Umfrage zum Technikeinsatz in Pflegeeinrichtungen (UTiP), Sachbericht, Bundesministerium für Gesundheit, Berlin, Germany, 2020. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/5_Publikationen/Pflege/Berichte/2020-06-26_IGES_UTiP_Sachbericht.pdf.

[5] C. Meske, B. Abedin, M. Klier, F. Rabhi, Explainable and responsible artificial intelligence, Electronic Markets 32 (2022) 2103–2106. URL: https://doi.org/10.1007/s12525-022-00607-2. doi:10.1007/s12525-022-00607-2.

[6] A. Adadi, M. Berrada, Peeking inside the black-box: A survey on explainable artificial intelligence (xai), IEEE Access 6 (2018) 52138–52160. doi:10.1109/ACCESS.2018.2870052.

[7] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, J. Zhu, Explainable ai: A brief survey on history, research areas, approaches and challenges, in: Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II, Springer-Verlag, Berlin, Heidelberg, 2019, p. 563–574. doi:10.1007/978-3-030-32236-6_51.

[8] D. Gunning, E. Vorm, J. Y. Wang, M. Turek, Darpa's explainable ai (xai) program: A retrospective, Applied AI Letters 2 (2021) e61. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ail2.61. doi:https://doi.org/10.1002/ail2.61. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/ail2.61.

[9] J. Hermann, N. Jansen, A. Dogangün, We need to understand where the data comes from: Co-designing transparent ai systems for caregiving, in: Proceedings of the 30th International Con-

ference on Intelligent User Interfaces Companion (IUI Companion ’25), Association for Computing Machinery, ACM, Cagliari, Italy, 2025. doi:10.1145/3708557.3716354.
[10] J. Hermann, A. D. Mäder, A.-K. Kubullek, C.-C. Hey, A. Dogangün, (intelligent) technical systems in elderly care: The caregivers perspective, in: Proceedings of the 34th Australian Conference on Human-Computer Interaction, OzCHI ’22, Association for Computing Machinery, New York, NY, USA, 2023, p. 81–87. URL: https://doi.org/10.1145/3572921.3572948. doi:10.1145/3572921.3572948.
[11] F. Bendig, E. Naroska, M. Weberskirch, An approach to privacy-aware image analysis on edge devices using CNNs, Current Directions in Biomedical Engineering 9 (2023) 262–265. URL: https://doi.org/10.1515/cdbme-2023-1066https://www.degruyter.com/document/doi/10.1515/cdbme-2023-1066/html. doi:10.1515/cdbme-2023-1066.
[12] P. Mohassel, Y. Zhang, Secureml: A system for scalable privacy-preserving machine learning, in: 2017 IEEE Symposium on Security and Privacy (SP), 2017, pp. 19–38. doi:10.1109/SP.2017.12.
[13] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models, ACM Comput. Surv. 51 (2018). URL: https://doi.org/10.1145/3236009. doi:10.1145/3236009.
[14] S. Gregor, I. Benbasat, Explanations from intelligent systems: theoretical foundations and implications for practice, MIS Q. 23 (1999) 497–530. URL: https://doi.org/10.2307/249487. doi:10.2307/249487.
[15] J. S. Christian Meske, Enrico Bunde, M. Gersch, Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities, Information Systems Management 39 (2022) 53–63. URL: https://doi.org/10.1080/10580530.2020.1849465. doi:10.1080/10580530.2020.1849465.
[16] G. Ras, M. van Gerven, P. Haselager, Explanation Methods in Deep Learning: Users, Values, Concerns and Challenges, Springer International Publishing, Cham, 2018, pp. 19–36. URL: https://doi.org/10.1007/978-3-319-98131-4_2. doi:10.1007/978-3-319-98131-4_2.
[17] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, L. Kagal, Explaining explanations: An overview of interpretability of machine learning, in: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 2018, pp. 80–89. doi:10.1109/DSAA.2018.00018.
[18] M. Li, S. Gregor, Outcomes of effective explanations: Empowering citizens through online advice, Decis. Support Syst. 52 (2011) 119–132. URL: https://doi.org/10.1016/j.dss.2011.06.001. doi:10.1016/j.dss.2011.06.001.
[19] J. Schneider, J. P. Handali, Personalized explanation for machine learning: A conceptualization, in: Proceedings of the 27th European Conference on Information Systems (ECIS), AIS, Stockholm & Uppsala, Sweden, 2019. URL: https://aisel.aisnet.org/ecis2019_rp/171.
[20] I. B. Jasbir S. Dhaliwal, The use and effects of knowledge-based system explanations: Theoretical foundations and a framework for empirical evaluation, Information Systems Research 7 (1996) 342–362. doi:10.1287/isre.7.3.342.
[21] I. B. Ji-Ye Mao, The Use of Explanations in Knowledge-Based Systems: Cognitive Perspectives and a Process-Tracing Analysis, Journal of Management Information Systems 17 (2000) 153–179. URL: https://doi.org/10.1080/07421222.2000.11045646. doi:10.1080/07421222.2000.11045646.
[22] R. Gozalo-Brizuela, E. C. Garrido-Merchán, Chatgpt is not all you need. a state of the art review of large generative ai models, Preprint, ArXiv abs/2301.04655, 2023.
[23] A. Leite-Moreira, A. Mendes, A. Pedrosa, A. Rocha-Sousa, A. Azevedo, A. Amaral-Gomes, C. Pinto, H. Figueira, N. R. Pereira, P. Mendes, T. Pimenta, An NLP solution to foster the use of information in electronic health records for efficiency in decision-making in hospital care, CoRR abs/2202.12159 (2022). URL: https://arxiv.org/abs/2202.12159. arXiv:2202.12159.
[24] A. E. W. Johnson, L. Bulgarelli, T. J. Pollard, Deidentification of free-text medical records using pre-trained bidirectional transformers, in: Proceedings of the ACM Conference on Health, Inference, and Learning, CHIL ’20, Association for Computing Machinery, New York, NY, USA, 2020, p. 214–221. URL: https://doi.org/10.1145/3368555.3384455. doi:10.1145/3368555.3384455.
[25] K. Holtzblatt, J. B. Wendell, S. Wood, Rapid contextual design: A how-to guide to key techniques

for user-centered design, Ubiquity 2005 (2005) 3. URL: https://doi.org/10.1145/1066348.1066325. doi:10.1145/1066348.1066325.
[26] B. Zhou, G. Yang, Z. Shi, S. Ma, Natural language processing for smart healthcare, IEEE Reviews in Biomedical Engineering 17 (2024) 4–18. doi:10.1109/RBME.2022.3210270.
[27] A. Bhattacharya, Applied Machine Learning Explainability Techniques: Make ML models explainable and trustworthy for practical applications using LIME, SHAP, and more, Packt Publishing, Birmingham, UK, 2022.
[28] M. T. Ribeiro, S. Singh, C. Guestrin, "why should i trust you?": Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 1135–1144. URL: https://doi.org/10.1145/2939672.2939778. doi:10.1145/2939672.2939778.
[29] M. T. Ribeiro, S. Singh, C. Guestrin, Anchors: high-precision model-agnostic explanations, in: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence, AAAI'18/IAAI'18/EAAI'18, AAAI Press, 2018.
[30] F. Gutiérrez, N. N. Htun, V. Vanden Abeele, R. De Croon, K. Verbert, Explaining call recommendations in nursing homes: a user-centered design approach for interacting with knowledge-based health decision support systems, in: Proceedings of the 27th International Conference on Intelligent User Interfaces, IUI '22, Association for Computing Machinery, New York, NY, USA, 2022, p. 162–172. URL: https://doi.org/10.1145/3490099.3511158. doi:10.1145/3490099.3511158.
[31] H. Yang, J. M. Garibaldi, Automatic detection of protected health information from clinic narratives, J. of Biomedical Informatics 58 (2015) S30–S38. URL: https://doi.org/10.1016/j.jbi.2015.06.015. doi:10.1016/j.jbi.2015.06.015.
[32] T. Miller, Explanation in artificial intelligence: Insights from the social sciences, Artificial Intelligence 267 (2019) 1–38. URL: https://www.sciencedirect.com/science/article/pii/S0004370218305988. doi:https://doi.org/10.1016/j.artint.2018.07.007.
[33] G. Vilone, L. Longo, Explainable artificial intelligence: A systematic review, Preprint, ArXiv abs/2006.00093, 2020. URL: https://arxiv.org/abs/2006.00093.
[34] B. Goodman, S. Flaxman, European union regulations on algorithmic decision making and a "right to explanation", AI Mag. 38 (2017) 50–57. URL: https://doi.org/10.1609/aimag.v38i3.2741. doi:10.1609/aimag.v38i3.2741.