How to Enable Sovereign Human-AI Interactions at Work? Concepts of Graspable Testbeds Empowering People to Understand and Competently Use AI-Systems

Wienrich, Carolin; Carolus, Astrid; Latoschik, Marc Erich

Erstveröffentlichung / Primary Publication
Konferenzbeitrag / conference paper

Empfohlene Zitierung / Suggested Citation:
Wienrich, C., Carolus, A., & Latoschik, M. E. (2023). How to Enable Sovereign Human-AI Interactions at Work? Concepts of Graspable Testbeds Empowering People to Understand and Competently Use AI-Systems. In B. Herlo, & D. Irrgang (Eds.), *Proceedings of the Weizenbaum Conference 2022: Practicing Sovereignty - Interventions for Open Digital Futures* (pp. 209-220). Berlin: Weizenbaum Institute for the Networked Society - The German Internet Institute. https://doi.org/10.34669/wi.cp/4.20

Proceedings of the Weizenbaum Conference 2022:
Practicing Sovereignty. Interventions for Open Digital Futures

HOW TO ENABLE SOVEREIGN HUMAN-AI INTERACTIONS AT WORK?

CONCEPTS OF GRASPABLE TESTBEDS EMPOWERING PEOPLE TO UNDERSTAND AND COMPETENTLY USE AI-SYSTEMS

Wienrich, Carolin
Julius-Maximilians-University
Wuerzburg, Germany
carolin.wienrich@uni-wuerzburg.de

Carolus, Astrid
Julius-Maximilians-University
Wuerzburg, Germany
astrid.carolus@uni-wuerzburg.de

Latoschik, Marc Erich
Julius-Maximilians-University
Wuerzburg, Germany
marc.latoschik@uni-wuerzburg.de

KEYWORDS

AI literacy; eXTended AI; human-centered AI; value-sensitive design; human-centered design

DOI: 10.34669/wi.cp/4.20

ABSTRACT

Artificial intelligence (AI) strategies are exhibiting a shift of perspectives, focusing more intensively on a more human-centric view. New conceptualizations of AI literacy (AIL) are being presented, summarizing the competencies human users need to successfully interact with AI-based systems. However, these conceptualizations lack practical relevance. In view of the rapid pace of technological development, this contribution addresses the urgent need to bridge the gap between theoretical concepts of AIL and practical requirements of working environments. It transfers current conceptualizations and new principles of a more human-centered perspective on AI into professional working environments. From a psychological perspective, the project focuses on emotional-motivational, eudaimonic, and social aspects. Methodologically, the project presented develops AI testbeds in virtual reality to realize literally graspable interactions with AI-based technologies in the actual work environment. Overall, the project aims to increase the competencies and the willingness to successfully master the challenges of the digitalized world of work.

1 INTRODUCTION

Digitalization is omnipresent in almost all areas of our lives. Developments in the field of machine learning and artificial intelligence (AI) are changing both our private and professional lives. Competencies and skills that allow for successful interactions with AI-based technologies are essential prerequisites for individuals to reap the societal benefits of digitalization (Law et al., 2018). The scope of public and scientific discussions of successful approaches of digitalization has long been limited to technical equipment and technical operating skills (Carolus & Wienrich, 2021; Wienrich & Carolus, 2021). Particularly in the context of work, the human has been unilaterally defined as *Homo economicus*, who regard technology as a tool to achieve goals, which are defined by the organizational framework and requirements (Carolus & Wienrich, 2021; Wienrich et al., 2022). More recent national and international AI strategies have argued for a more human-centric transformation process focusing on the human being and the individual's cognitive, emotional, and conative processes. For example, the German governmental AI strategy summarizes critical points of AI-related transformation processes as follows: "It is about individual freedom rights, autonomy, personal rights, the individual's freedom of choice. About hopes, fears, potentials, and expectations" (Bundesministerium für Bildung und Forschung, n.d.). In this context, the European Union has made considerable efforts to draft viable rules for the new world of AI. Within this framework, the EU High-Level Expert Group on AI established requirements such as: AI must be trustworthy (e.g., technically and socially robust) and respect human-centric approaches (e.g., respect human needs and diversity, avoid discrimination, be explainable). Thus, the previously merely technical focus was expanded to include an explicitly human-centric perspective. In particular, Article 14 of the AI regulation (Lexparency.org, n.d.) and the corresponding statements in the white AI paper (Madiega, 2022) emphasize that human needs should be taken into account in terms of both the design and the use of AI systems. Transparency and participation are regarded as essential elements to ensure trust in AI applications.

The present contribution emphasizes the urgency of translating these crucial but still rather theoretical demands into concrete professional practice. Furthermore, it bridges another gap: large parts of both public debates and governmental regulations refer to the business practices of only a few large high-tech companies. The reality of smaller and medium-sized companies and their employees is, however, neglected. The present project addresses these desiderata and focuses on people who work or will (soon) work with digital entities and AI-based systems—people who experience considerable digital inequalities and disparities. Workplaces are thereby considered to represent more to human beings than mere professional working spheres. Instead, workplaces are

essential social spaces that fulfill various human needs above and beyond professional efficiency and effectivity. Consequently, change processes at work affect employees at different levels, depending on the underlying basic human needs and interindividual differences: prior experience with technology and computer science, dystopian fears about AI or the individual's working environment and degree of participation, and digital disadvantages.

2 SOVEREIGNTY AND AI LITERACY AT WORK: RELATED WORK AND DESIDERATA

The concepts of media literacy and, more recently, digital literacy have been widely discussed as (rather) new cultural techniques that are an essential prerequisite for successful participation in our present and future digitized world. With the increasing importance of AI, these concepts of technology-related competencies need to be updated. The recently introduced concept of AI literacy (short: AIL) aims at understanding the competencies enabling people to successfully interact with AI-based technologies.

2.1 DESIDERATUM 1: THE NEED FOR VALID AND STRAIGHTFORWARD MEASURES OF AI LITERACY

The acquisition of digital (AI) literacy is crucial in order to competently harness future-oriented technologies. In a professional environment, there is a growing demand for new competency profiles that are increasingly characterized by associated AI-related innovations. Long and Magerko (2020) have analyzed competencies that enable the individual to comprehend, critically evaluate, and competently use AI technologies. The authors present a competency grid consisting of 17 AI-relevant skills, unified under the umbrella term *artificial intelligence literacy* (AIL; Long & Magerko, 2020). In an expert workshop, Wienrich et al. (2022), together with the AI Observatory of the German Federal Ministry of Labor and Social Affairs, expanded this framework. With an explicit focus on the work environment, their *Competence Behavioral Model of AI Literacy* (CBM-AIL) embeds individual and organizational potentials and barriers on the micro-, meso-, and macro-level. Besides these nascent conceptualizations, there are very limited measuring approaches that target general digital literacy (e.g., Jenkins et al., 2006; Ng, 2012; Ng et al., 2021; Porat et al., 2018) but not AIL (Wienrich & Carolus, 2021). Moreover, these measures are rather extensive questionnaires that are often unsuitable for practitioners or refer only to voice-based AI systems. Further studies analyzed perceptions and associations in the context of AI (e.g., European Commission, 2017; Hadan & Patil, 2020; Lau et al., 2018; Kelley et al., 2019; Zeng et al., 2017; Zhang & Dafoe, 2019). However, these studies often used measures consisting of only a single item, resulting in rather vague latent variables and limited validity. In summary, existing conceptualizations of AI-related competencies lack

practical and professional relevance. Additionally, the quality of existing measures is limited in terms of the scientific, practical, and work-related criteria.

2.2 Desideratum 2: The Need for Holistic Human-Centered Conceptualizations

Digital transformation promotes methods and disciplines that provide insights into either human processes (e.g., psychology) or human-centered approaches (e.g., human-computer interaction, HCI). The keywords explainable AI (e.g., Goebel et al., 2018; Samek et al., 2019) or human-centered AI (e.g., Riedl, 2019; Xu, 2019) refer to ongoing efforts to increase explainability and intuitive usability in the context of both the development and the implementation of AI systems (e.g., Wittpahl, 2019; Kraus et al., 2021). Approaches so far (at least implicitly) tend to follow the idea of the *Homo Economicus* and are therefore often limited to the analysis of pragmatic qualities (usability), while hedonic, eudaimonic, and social aspects—as well as professional and work-related aspects—are neglected (Carolus & Wienrich, 2021; Wienrich et al., 2022). Thus, this conception of the human limits the understanding of human-technology interactions as it neglects psychological and HCI knowledge. Only when this knowledge is taken into account do we adequately reflect the effects of human experiences, actions, and feelings. Consequently, skills and metacognitive competencies can be regarded as equally relevant as feelings of meaning and self-actualization (e.g., Mekler & Hornbæk, 2016).

AI-based technologies change the conceptualization of technology in general. The increasingly intelligent and interactive functions of AI turn what were once digital tools into (social) counterparts and interaction partners (Carolus et al., 2019; for a recent overview: Li & Suh, 2021; see also Reeves & Nass, 1996; Carolus et al., 2021; Wienrich et al., 2021).

The wider concept of user experience, which is enriched with emotions, perceptions, preferences, physiological and psychological reactions, behaviors, and performances occurring before, during, and after the interaction with technology, is regarded as essentially important for the consideration of usability in the professional work environment (Bargas-Avila & Hornbæk, 2011; Hassenzahl et al., 2010; Pataki et al., 2006). However, the current state of research reveals desiderata in terms of both the more comprehensive conceptualization of user experience and a widely accepted set of holistic and human-centered criteria for the evaluation of AI-based systems in the professional work environment.

2.3 Desideratum 3: The Need for Self-Determination

Comprehensive technological developments increase the complexity of systematic and valid investigations of human-technology interactions. Wienrich and Latoschik (2021) introduce the

eXtended AI approach as a new research heuristic that uses extended realities (XR; e.g., virtual reality, augmented reality) to enable systematic and valid investigations of AI systems and their effects. The basic idea is to use XR technology to create innovative testbeds reflecting complex AI interfaces and human-AI interactions. Utilizing rapid prototyping methods, *eXtended AI* reduces complexity and provides immersive experiences into certain fields of application. Moreover, various design spaces, easy accessibility, versatility, and tangible training possibilities are promising benefits of the *eXTended AI* approach. Possible forms of human-robot interaction (for instance, in logistics workplaces) were simulated virtually in a pilot study. Results show that different design features of the virtual robot contributed differently to fulfilling the diverse human needs at work. Significant gender differences were also identified. So far, the eXtended AI approach has only been studied in laboratory experiments with student samples—but not in professional work scenarios. As mentioned, there is a particular need for action in these professional contexts: While technologies and user requirements change rapidly, (future) users are rarely involved in change processes at an early stage. Approaches that allow low-threshold access to future technologies are still rare. This exclusion of users is highly problematic, as recent occupational psychology studies emphasize that limited opportunities to participate in the planning process lead to higher levels of dissatisfaction and increased workload (Carls et al., 2021). Conversely, the early involvement of users in the digital transformation process will positively affect the employees' work satisfaction.

To summarize, this study identifies a need for new participatory and human-centered approaches to introduce and implement AI in the workplace. These new approaches can contribute to the participatory, human-centered derivation of implementation requirements.

3 Concepts of Grasping AI Testbeds for Empowering Workers to Use and Competently Use AI

3.1 Addressing Desideratum 1: Design VALID and Straightforward Measures of AI Literacy

On the one hand, existing measures of AI literacy were mostly developed in the laboratory or under controlled conditions and with student samples who received course credit for participation. This resulted in large quantities of items and scales in a rather academic language. On the other hand, single-item measures and further less valid and reliable assessments are often used in studies in the field. Addressing both shortcomings, this study proposes a new strategy for developing practical measurements. In a top-down process, items and scales from existing scientific models are derived to create a scientifically reliable item pool. In a bottom-up process, the perspectives of domain experts and lay people who are affected by the introduction of AI systems at work are mapped to the item

pool. The resulting items are clustered into small modules and then adapted and reduced. The early involvement of people who are affected by technological change processes meets one of the essential requirements of various approaches (e.g., user-centered design, contextual design, value-sensitive design). Furthermore, the development of the measure of AI literacy is based on fundamental and applied research. The overall goal: short, comprehensive, valid, and reliable measures for AIL that fulfill both scientific (quality criteria) and practical requirements (increase the individuals' willingness to participate).

3.2 Addressing Desideratum 2: A Holistic Perspective on Human Beings at Work

As mentioned above, workspaces are highly associated with various human needs and motives. Consequently, the psychological perspective of this project emphasizes the importance of emotional and motivational aspects at work. For example, efficiency and effectiveness are equally important for the feeling of acceptance and community. Following this perspective, the project follows a holistic concept of the human being at work (going beyond the idea of the *Homo oeconomicus*). It considers emotional-motivational aspects as well as hedonic, eudaimonic, and social motives and needs. Moreover, the project will analyze the attribution of human characteristics to technology, which shape the users' expectations and interaction behaviors. Finally, the project considers the ongoing shift in perspectives on technology—from a perspective that considered technology as a tool to one that focuses on intelligent, interactive digital entities. Taken together, these three aspects will determine the criteria for requirements analyses, the design and development processes, and the practical implementations.

3.3 Addressing Desideratum 3: eXtended AI at Work

In everyday business, most of the employees who use technology are not involved in its development and implementation. However, excluding them at the early stage is risky. Mismatches between the systems and the employees' needs, which are detected after the system has been incorporated into the organizational processes, are cost intensive. Therefore, on the one hand, employee participation is promising. On the other hand, their perspective is limited in terms of their power of imagination and the validity of their projections. To close this gap, the present project proposes "graspable interactions" using the vast potential of XR. XR-based applications allow cost-efficient prototyping, which may vary in the degree of complexity and realism. Thus, XR is regarded as a powerful design space that can embody AI-based interaction partners without being bound to the limitations of conventional prototypes (e.g., physical environment, engineering challenges, costs). Furthermore, XR allows researchers to study heterogeneous user groups by adapting the prototypes to the individual's

degree of expertise or to different domains and tasks. Finally, XR provides a safe testbed without actual consequences in the real world (e.g., damages).

In sum, the project considers XR as a promising way to involve employees by providing AI experiences and explaining and illustrating the functionality and the consequences of the technology during the actual interaction. They can learn about AI, how to use and how to adapt the systems to their needs. Companies can trigger curiosity instead of fear and distrust by promoting participation and human-centeredness, development, and implementation.

4 Practical Examples Across Different Workplaces

4.1 Example 1: Robots in Logistics

Industrial robots are already used in various ways; they mostly work autonomously and at a safe distance from humans. In the future, robots are likely to collaborate more directly with humans (e.g., handing over work pieces). Hence, the design and behavior of robots (e.g., indicating social behavior or social intelligence) will likely have an impact on the acceptance and the human-robot work performance. XR testbeds can simulate the workplace and the robot in multiple versions. Employees can test and evaluate the different prototypes to contribute to further development—not only in line with economic and technical requirements but also in line with their human needs.

4.2 Example 2: Recommender Systems in Administration

AI can support administrative tasks in many ways. Against the background of the current legal situation, however, the human being is ultimately responsible for the decision. Therefore, the question arises of how AI can support human employees by providing information that is both relevant and accepted. Since human decisions are subject to numerous psychological biases, it is necessary to consider, for example, which interface and which information presentation lead to the best results, hence meeting the requirements of both the best decision and the human needs. In XR, different recommender systems can be tested and evaluated to contribute to further development. In addition to the administration, other domains are conceivable (e.g., medicine, law, insurance).

4.3 Example 3: Speech-Based Systems in Customer Service

Chatbots and voice assistants provide increasing support in customer service, mostly in addition to human consultants. Research shows that the design of the entities has a significant impact on the perception, acceptance, and trust of the systems and the service. In XR, different systems with

different designs and outward appearances can be tested and evaluated to contribute to further development.

5 Conclusion and Contribution

This paper emphasizes the urgency of transferring theoretical knowledge about critical demands for the development and implementation of human-centered AI systems into professional practice. While current public and political discussions tend to be limited to the business processes of only a few high-tech global players, this project focuses on people who work or will work with digital entities and AI systems. Moreover, the project incorporates a psychological perspective on increasingly digitalized workplaces, which are defined as essentially social spaces associated with multiple basic human needs. Thus, the employees' perspective—going beyond effectiveness and efficiency—becomes more complex and more important. Digital change processes of workplaces affect human beings, who are driven by their human needs. Additionally, employees come from different backgrounds and are equipped with different prior knowledge in computer science, resulting in interindividual differences in terms of their expectations from AI (dystopian fears vs. overestimating benefits), their participation and involvement at work, and their feelings of digital inequalities and disparities.

Summarizing the status quo, the project focuses on three desiderata: (1) the lack of valid, reliable, and practical measures of AIL, (2) untapped potential that arises from a holistic view of human-technology interaction that integrates human information processing and need structures, and (3) benefits of participative processes in development and implementation. These desiderata strongly contradict the demands of national and international AI strategies and endanger the safety of people and technology, job satisfaction and commitment, and value creation. To counteract this, our project proposes new strategies for socio-technical education and pedagogy in the context of work. It translates the critical demands and theoretical considerations from both science and politics into professional practice following the demands of the German federal government to contribute to the various layers of a socially responsible digital transformation.

6 Acknowledgments

This research has been funded by the German Federal Ministry of Labor and Social Affairs in the project AIL AT WORK (project number DKI.00.00030.21).

7 REFERENCES

1. Bargas-Avila, J. A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: A critical analysis of empirical studies of user experience [Conference paper]. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/1978942.1979336
2. Bundesministerium für Bildung und Forschung. (n.d.). *Nationale Strategie für Künstliche Intelligenz* [National strategy for artificial intelligence]. https://www.ki-strategie-deutschland.de/home.html
3. Carls, K., Gehrken, H., Kuhlmann, M., Thamm, L., & Splett, B. (2021). Digitalisierung, Arbeit und Gesundheit [Digitalisation, work and health]. In K-P. Buss, M. Kuhlmann, M. Weißmann, H. Wolf & B. Apitzsch (Eds.), *Digitalisierung und Arbeit – Triebkräfte, Arbeitsfolgen, Regulierung* (pp. 235-273). Campus.
4. Carolus, A., Binder, J. F., Muench, R., Schmidt, C., Schneider, F., & Buglass, S. L. (2019). Smartphones as digital companions: Characterizing the relationship between users and their phones. *New Media & Society*, *21*(4), 914-938. https://doi.org/10.1177/1461444818817074
5. Carolus, A., & Wienrich, C. (2021). Towards a holistic approach and measurement of humans interacting with speech-based technology. In C. Carolus, C. Wienrich & I. Siebert (Eds.), *1st AI-debate workshop: Workshop establishing an interdisciplinary perspective on speech-based technology* (pp. 39-41). Otto von Guericke University Magdeburg. http://dx.doi.org/10.25673/38471
6. Carolus, A., Wienrich, C., Törke, A., Friedel, T., Schwietering, C., & Sperzel, M. (2021). 'Alexa, I feel for you!'-Observers' empathetic reactions towards a conversational agent. *Frontiers in Computer Science*, *3*, Article e682982. https://doi.org/10.3389/fcomp.2021.682982
7. European Commission. (2017). *Special Eurobarometer 460: Attitudes towards the impact of digitisation and automation on daily life*. https://ec.europa.eu/jrc/communities/sites/jrccties/files/ebs_460_en.pdf
8. Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P., & Holzinger, A. (2018). Explainable AI: The new 42? In A. Holzinger, P. Kieseberg, A. Min Tjoa & E. Weippl (Eds.), *Lecture notes in computer science: Vol. 11015. Machine learning and knowledge extraction* (pp. 295-303). Springer. https://doi.org/10.1007/978-3-319-99740-7_21
9. Hadan, H., & Patil, S. (2020). Understanding perceptions of smart devices. In M. Bernhard, A. Bracciali, L. J. Camp, S. Matsuo, A. Maurushat, P. B. Rønne & M. Sala (Eds.), *Lecture notes in computer science: Vol. 12063. Financial cryptography and data security* (pp. 102-121). Springer. https://doi.org/10.1007/978-3-030-54455-3_8
10. Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, and interactive products – Facets of user experience. *Interacting with Computers 22*(5), 353–362. https://doi.org/10.1016/j.intcom.2010.04.002
11. Jenkins, H., Clinton, K., Purushotma, R., Robison, A. J. & Weigel, M. (2006). *Confronting the challenges of participatory culture: Media education for the 21st century*. The MacArthur Foundation. https://files.eric.ed.gov/fulltext/ED536086.pdf
12. Kelley, P. G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T., & Woodruff A. (2019). Exciting, useful, worrying, futuristic: Public perception of artificial intelligence in 8 countries. In M. Fourcade, B. Kuipers, S. Lazar, & D. Mulligan (Eds.), *Proceedings of the 2021 AAAI/ACM conference on AI, ethics, and society* (pp. 627–637). Association for Computing Machinery. https://arxiv.org/abs/2001.00081

13. Kraus, O., Ganschow, L., Eisenträger, M., & Wischmann, S. (2021). *Erklärbare KI. Anforderungen, Anwendungsfälle und Lösungen* [Explainable AI requirements, use cases and solutions]. Bundesministeriums für Wirtschaft und Energie. https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/KI-Inno/2021/Studie_Erklaerbare_KI.pdf;jsessionid=C94CE1AF540F11B6D64E3EB19A61A2AF?__blob=publicationFile&v=9

14. Lau, J., Zimmerman, B., & Schaub, F. (2018). Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction, 2*(CSCW), Article 102. https://doi.org/10.1145/3274371

15. Law, N., Woo, D., de la Torre, J., & Wong, G. (2018). A global framework of reference on digital literacy skills for indicator 4.4.2 (Information paper No. 51; p. 146). UNESCO Institute for Statistics. http://uis.unesco.org/sites/default/files/documents/ip51-global-framework-reference-digital-literacy-skills-2018-en.pdf

16. Lexparency.org. (n.d.). *Art. 14 AI-regulation (proposal) - Human oversight.* Retrieved September 1, 2022, from https://lexparency.org/eu/52021PC0206/ART_14/

17. Li, M., & Suh, A. (2021). Machinelike or humanlike? A literature review of anthropomorphism in AI-enabled technology. In T. X. Bui (Ed.), *Proceedings of the 54th Hawaii International Conference on system sciences* (pp. 4053-4062). Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2021.493

18. Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations [Conference paper]. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3313831.3376727

19. Madiega, T. (2022). *Artificial intelligence act*. European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf

20. Mekler, E. D. & Hornbæk, K. (2016). Momentary pleasure or lasting meaning? Distinguishing eudaimonic and hedonic user experiences [Conference paper]. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. http://dx.doi.org/10.1145/2858036.2858225

21. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article e100041. https://doi.org/10.1016/j.caeai.2021.100041

22. Ng, W. (2012). Can we teach digital natives digital literacy? *Computer & Education, 59*(3), 1065–1078. https://doi.org/10.1016/j.compedu.2012.04.016

23. Pataki, K., Sachse, K., Prümper, J., & Thüring, M. (2006). ISONORM 9241/10-S: Kurzfragebogen zur Software-Evaluation [ISONORM 9241/10-S: Short questionnaire for software evaluation]. In F. Lösel & D. Bender (Eds.), *Kongress der Deutschen Gesellschaft für Psychologie (45.) – Humane Zukunft gestalten* (pp. 258-259). Pabst Science Publishers.

24. Porat, E., Blau, I., & Barak, A. (2018). Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. *Computer & Education, 126*, 23–36. https://doi.org/10.1016/j.compedu.2018.06.030

25. Reeves, B. & Nass, C. (1996). The media equation: How people treat computers, television, and new media like real people. Cambridge University Press

26. Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies, 1*(1), 33-36. https://doi.org/10.1002/hbe2.117

27. Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K. R. (Eds.). (2019). *Lecture notes in computer science: Vol. 11700. Explainable AI: interpreting, explaining and visualizing deep learning*. Springer Cham. https://doi.org/10.1007/978-3-030-28954-6

28. Wienrich, C., & Carolus, A. (2021). Development of an instrument to measure conceptualizations and competencies about conversational agents on the example of smart speakers. *Frontiers in Computer Science, 3*, Article e685277. https://doi.org/10.3389/fcomp.2021.685277

29. Wienrich, C., Carolus, A., Augustin, Y., & Markus, A. (2022). *AI Literacy: Kompetenzdimensionen und Einflussfaktoren im Kontext von Arbeit* [Working paper] [AI Literacy: Competence dimensions and influencing factors in the context of work]. KI-Observatorium des Bundesministeriums für Arbeit und Soziales. https://www.denkfabrik-bmas.de/fileadmin/Downloads/Publikationen/AI_Literacy_Kompetenzdimensionen_und_Einflussfaktoren_im_Kontext_von_Arbeit.pdf

30. Wienrich, C., & Latoschik, M. E. (2021). Extended artificial intelligence: New prospects of human-AI interaction research. *Frontiers in Virtual Reality, 2*, Article e686783. https://doi.org/10.3389/frvir.2021.686783

31. Wienrich, C., Reitelbach, C., & Carolus, A. (2021). The trustworthiness of voice assistants in the context of healthcare investigating the effect of perceived expertise on the trustworthiness of voice assistants, providers, data receivers, and automatic speech recognition. *Frontiers in Computer Science, 3*, Article e685250. https://doi.org/10.3389/fcomp.2021.685250

32. Wittpahl, V. (2019). *Künstliche Intelligenz: Technologien | Anwendung | Gesellschaft* [Artificial intelligence: Technologies | application | society]. Springer. https://doi.org/10.1007/978-3-662-58042-4

33. Xu, W. (2019). Toward human-centered AI: A perspective from human-computer interaction. *Interactions, 26*(4), 42-46. https://doi.org/10.1145/3328485

34. Zeng, E., Mare, S., & Roesner, F. (2017). End user security and privacy concerns with smart homes. In M. E. Zurko, S. Chiasson & M. Smith (Eds.), *Proceedings of the thirteenth symposium on usable privacy and security* (pp. 65-80). USENIX Association. https://www.usenix.org/system/files/conference/soups2017/soups2017-zeng.pdf

35. Zhang, B., & Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends*. Center for the Governance of AI, Future of Humanity Institute, University of Oxford. http://dx.doi.org/10.2139/ssrn.3312874