Human-Computer Interaction & Emerging Technologies, Vol. 233, 2026, 291–299
https://doi.org/10.54941/ahfe1007528

Towards Human-Centered Development of Explainable AI in a German SME

Nadja Hemming, Simeon Ackermann, and Amit Kirschenbaum

Competence Center KMI at Institute for Applied Informatics (InfAI e.V.), 04109 Leipzig, Germany

ABSTRACT

Small and medium-sized enterprises (SMEs) in the manufacturing and maintenance industry increasingly adopt AI systems in a variety of applications. Therefore, the design and implementation of software environments in SMEs must meet specific requirements, such as human-centered design (HCD) and explainability of AI (XAI). Yet, applied research that utilizes HCD for XAI is sparse. We aim to implement an AI system assigning employees to specific tasks in a maintenance SME. Our work-in-progress report examines an approach for an agile human-centered implementation of an XAI system, including challenges and benefits. We expect to improve the data quality and AI predictions on the one hand. On the other hand, we aim to improve the acceptance of the use of AI and the organizational work process. With further investigations, we deepen the discussion on feasibility of human-centered development of XAI in this context.

Keywords: Explainable AI, Human-centered design, Small-and-medium-sized enterprises, Competence matching

INTRODUCTION

The rapid and dynamic adoption of artificial intelligence (AI) has witnessed an increased interest and application in most domains such as healthcare, education, law, finance, and various industrial contexts. Research on AI adoption into widespread industrial practice, especially among small and medium-sized enterprises (SMEs), remains underrepresented in comparison to already mentioned high-stake domains (Bunte et al., 2021). AI adoption as a specific SME challenge stresses among other aspects the growing importance of trust-promoting factors such as explainability.

To illustrate this, the examined SME consists of approximately 150 employees at five different locations surrounding a town in the federal state of Saxony in Germany. Our use case considers staff assignment, hereby referring to the process of allocating personnel to specific projects based on their skills, availability, and work priority. The company under discussion operates in the machine construction, maintenance, and repair sector. In daily business, it receives service calls from customers, and maintenance coordinators are responsible for the allocation of personnel, so-called maintenance technicians.

The objective of the presented use case is to provide AI-supported recommendations for personnel planning. Based on task descriptions,

Received February 6, 2026; Revised March 30, 2026; Accepted April 13, 2026; Available online July 20, 2026

appropriate competencies are assigned using the machine learning model decision tree (DT).

Studies addressing similar topics (Yingbo et al., 2007; Mo et al., 2020) view this merely as a machine learning problem and do not involve any relevant companies to get a practical orientation to the topic during the development phase. They also fail to address human perspectives of the process, e.g. to acquaint themselves with the needs and requirements of the personnel and clients.

This paper is structured as follows: Section 2 examines both the SME context and underlying core concepts. Section 3 introduces the research questions and section 4 establishes a description of the methodology. Section 5 introduces preliminary results. Finally, the conclusion discusses future implications and the next steps.

THEORETICAL BACKGROUND

SMEs in Industrial Contexts

SMEs among the German manufacturing industry increasingly adopt AI systems in various applications (Schwaeke et al., 2025), although findings showcase that in comparison to other countries there is more concern about core digitalization measures and preliminary steps prior to AI adoption (Merkel-Kiss and Von Garrel, 2022; Hein et al., 2025). SMEs operate under constraints, such as limited human and financial resources, and heterogenous technological skills. SMEs face pronounced human resource constraints, e.g. a shortage of AI specialists (Schwaeke et al., 2025) and multiple organizational roles within a single employee. Financial constraints restrict SMEs' ability to invest both in AI technologies and infrastructure (OECD, 2025).

Human-Centered Design (HCD) and Agile Development

Human-centered design (HCD) can be summarized as a design approach that incorporates people with different prior experiences, desires, needs, ambitions, interests, irrational decision making, and lifestyles situated in specific cultural contexts in the development of a design process (Auernhammer, 2020). HCD is a shift of viewing humans as central in every aspect of the design. To our understanding, HCD emphasizes involving various users and indirectly affected groups as active contributors to the development process (Schmager et al., 2025).

We are guided by the human–centered AI (HCAI) approach postulated by Shneiderman (2020) that encompasses the possibility of achieving both high automation levels and a high level of human oversight (Campos and Shahkovska, 2025). Additionally, Friedrich et al. (2024) emphasize a clear and actionable guideline for navigating the complexities of AI-driven change for SMEs.

We follow HCD development based on agile principles such as sprint iterations, continuous improvement, and feedback loops. Built on these principles, we want to ensure that we integrate a structured, interdisciplinary and transparent methodology as a robust foundation for our iterative refinements, with a human at the center of the AI lifecycle (see Figure 1, Ozmen Garibay et al., 2023; Friedrich et al., 2024; Akthar et al., 2024).

Towards Explainability Embedded in a HCD Process

Explainable AI (XAI) is a subfield of AI that produces a set of algorithmic methods to not only make output comprehensible but also actionable for humans (Liao and Varshney, 2021). The aspect of explainability in this case is regarded as a human-centric characteristic as it is crucial for humans to trust, adopt, and effectively interact with real-world AI systems. Besides that, it is emphasized that explainability must be grounded in human-computer interaction and user experience design, hence current XAI research mostly focussed on mostly algorithmic interpretability, less on usability and interdisciplinary oversight (Liao and Varshney, 2021) as core human-centric aspects.

One key distinction between explainability approaches is the difference between post-hoc and ante-hoc methods: Post-hoc methods explain opaque ("black-box") models after they have been trained, to understand how they arrive to specific predictions by approximating their behavior, In contrast, ante-hoc models are characterized by their inherent transparency ("glass-box") and interpretability, which is intrinsic to their design, meaning that the logic behind their decision-making processes is evident in their structural design (Dwivedi et al., 2023). Decision tree models are of the latter type and have a long tradition since the beginning of AI (Holzinger, 2018). This intrinsic feature of decision trees may promote the trust of the decision and thus the acceptance of the AI usage.

RESEARCH INTEREST

We aim to expand applied research based on HCD and an agile development process for XAI within this industrial context. Our aim is to investigate the application of these approaches and discuss the particularities of the SME context. Explainability, usability, and interdisciplinary oversight will be included as core aspects in our approach. Interdisciplinarity can be guaranteed by consulting expertise of data science, software development, and human factors psychology. Based on this, we propose the following research questions:

RQ1: To what extent can human-centered XAI development be applied in this use case?
RQ2: How can a decision tree contribute to optimal XAI in this context?
RQ3: Which challenges emerge? Which countermeasures can be proposed in this specific SME context?

METHODOLOGY

We conceptualized our approach, depicted in Figure 1: the development process was separated into multiple iterations with different phases and activities. In contrast to traditional user-centered design, the involvement of all humans and stakeholders will be ensured by embedding them in all development stages and activities. Below, we look at the development divided into technical and human-centric aspects.

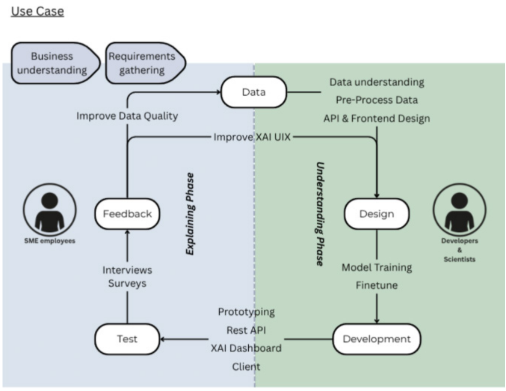

Figure 1: Human-centered agile framework for AI-XAI development. (Adapted from Campos and Shakhovska, 2025; Dwivedi et al., 2023).

Technical Aspects

During implementation, a key focus was to remain transparent. The technical components in the current project, mainly located in the understanding phase of Figure 1, can be divided into four parts:

1. Data pre-processing and model training
2. Prediction API development
3. Client development
4. XAI Dashboard development

Initially, one of the main challenges within this use case was data curation and data quality. The pre-processing of relevant data presented a profound challenge, as data quality and file structures varied and impacted the accuracy of predictions. At the beginning, the order descriptions were very short and contained ambiguous abbreviations which led to incorrect predictions. Therefore, we improved the data quality in several iterations in collaboration with the SME.

Another question arising during the understanding phase, concerned the choice of the model and deployment environment (self-hosted vs. cloud-provider). The following criteria were considered for decision: control vs. effort, data privacy, costs, trust, and acceptance. In our case, the system processes customer and order data, so for data privacy reasons, this precludes sharing data with a provider. Furthermore, the effort involved in operating an AI system themselves in-house must be weighed against the expected dependencies and service costs of the LLM providers, who sometimes offer a significantly broader range of functionality than a self-hosted solution.

The choice of a model was based on various criteria, gathered with stakeholders of the SME in the initial phase of the project, summarized in Table 1. A primary requirement was the system transparency and traceability to address potential conflicts in personnel planning, thus ensuring the medium-term acceptance of all stakeholders. The available resources (financial and personnel) have further limited development and hardware capacities. The main challenge of generating predictions from the descriptions can be classified as a so-called classification problem, in our case with unstructured text data.

Table 1: Characteristics of use case.

Aspect	Use Case SME Aspect
goal	Predict competences (text classification)
type of data	unstructured (short) text, semi-structured data
budget	low-cost demanded
complexity	low complexity but high classification challenge (e.g. because of data quality, ambiguity, text-shortness)
transparency	explainability demanded
accuracy	high accuracy improves employee acceptance and trust
privacy concerns	high, handling personal and company data
performance / inference time	acceptable time to generate answers is <5s

For the implementation of the prediction API, a decision tree model was chosen, because it is a robust predictive model in supervised machine learning, known for their unquestionable utility in a wide range of applications (Costa and Pedreira, 2023). It is suitable for solving non-binary (text) classification problems and can be trained and operated with low hardware resources. The DT model was trained across several development cycles and the most suitable models were evaluated in multiple experiments to define the text classifier algorithm (bag-of-words).

In collaboration with the SME and multiple iterations, a client was developed specifically tailored to integrate into the company's software. The client's primary function is to send the entered job description to the Prediction API and display the suggestions, the probability of relevance, and suitable maintenance technicians. The process is roughly illustrated in Figure 2, although the processes for accessing the XAI dashboard and reporting feedback on the current status are not yet fully implemented. Autonomy over personnel decisions (suggesting and selecting an employee) remains with the coordinators. A key non-functional requirement for the client was a lightweight and simple user interface and an accessible and intuitive user experience.

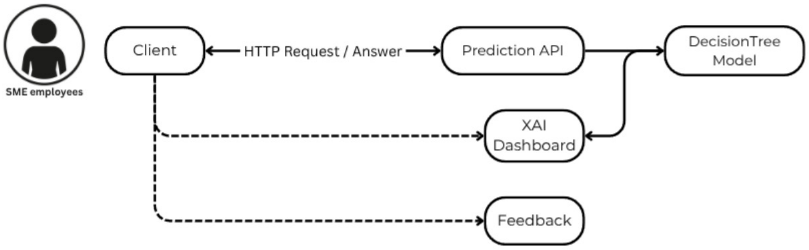

Figure 2: Client – Service interaction.

A key focus of the use case is to integrate transparency, traceability, and explainability of the AI decisions for all stakeholders, primarily of the SME's employees. To this end, an XAI dashboard was initially developed to illustrate the predictions from various perspectives, primarily from a scientific point of view (Figure 3). This includes e.g. the normalization of the order description and the impact of keywords on the prediction. It also highlights a bias that arises when unbalanced data distribution is not addressed.

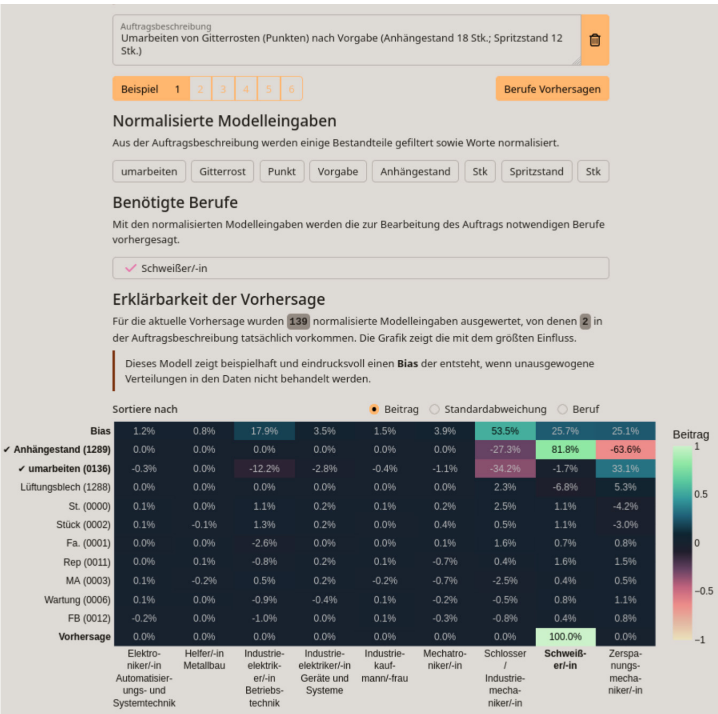

Figure 3: XAI dashboard, as of Jan 2026.

Human-Centric Aspects

Most employees working in maintenance will not interact with the AI system but instead be affected by AI-supported coordinator decisions. Moreover, the customers of the company will be indirectly affected as they need suitable and correct technician matches. The coordinators at various locations and the CEO will be in direct interaction with the AI system. Thus, a differentiation of various user groups and affected groups was established. Along the whole development cycle (see Figure 1), all these groups need to be asked for feedback. From an early stage on, we established direct and regular communication between the research team and the SME.

Primarily, to tailor to the needs of the company, the collaboration must be designed next to daily business, hence any communication needs to be precise, with everyone involved and several appointments to choose from need to be given to deal with shifting priorities. As a good practice it was successful to mix online and in-person regular meetings. The satisfaction of the company was reportedly high as they used the scientific input.

Secondly, for an early stage XAI dashboard iteration in the exploration phase, we identified the coordinator lead at one location as well as the company CEO as important domain-specific experts due to their organizational roles as both user and decision-maker. As part of an evaluation iteration in December 2025, we conducted two 45 min-lasting semi-structured interviews as part of the testing and feedback activities (as depicted in Figure 1). We included context-aware open questions such as "*How do you envision the client in an ideal world in order to ensure the best possible usability?*" or "*How can the XAI dashboard be designed to improve the explainability of the AI proposal?*". Next, we included Likert scale items on satisfaction, importance and convenience/usability of both the client and the XAI dashboard. With this in mind, we can quantitatively monitor these self-described aspects and retrace developments during the process. Finally, we presented and discussed three different prototypes of an XAI dashboard integrated into the client and evaluated the usability for further improvements.

PRELIMINARY RESULTS

We have analysed preliminary results obtained from the current technical implementation and HCD. SME specific requirements were gathered, and from these, the system environment as well as suitable algorithms and models were selected.

We have seen close and continuous cooperation as a feasible approach for the integration and acceptance of AI solutions. For this reason and to refer to the first research question, we rate the approach of HCD to develop XAI in this use case as feasible and estimate its impact as quite high due to feedback from the company (both interviewees rated the satisfaction and usability as very satisfying), domain-specific knowledge included at an early development stage and optimal resource management.

A DT model has been trained and implemented as a prediction REST-API service. In close collaboration with the SME, the data quality was continuously improved to further enhance the API's predictions. The XAI dashboard helped by identifying data bias as well as problematic abbreviations and too short and ambiguous descriptions thus addressing the research question 2. But still, providing a clear explanation in an XAI dashboard remains a challenge, and research question 3 will be answered in further studies.

FUTURE IMPLICATIONS AND CONCLUSION

On the technical side, we expect to continuously improve the quality of data and AI predictions. We assume that DTs will play a pivotal role in the future as there will be increasing opportunities to implement automated decision processes into day-to-day life, especially those in which interpretability is a key concern (Costa and Pedreira, 2023). Though, visualizing DTs in an understandable and helpful way for affected users remains a challenge.

On the human-centric side, we see the potential to improve the acceptance of AI adoption in the organization. Looking ahead, after the XAI dashboard reaches a fully matured and operational state, we will administer an online context-aware survey to other coordinators, including items from the system causability scale (Holzinger et al., 2023) to measure the quality of the explanations. Moreover, it will be beneficial to assess attitudes and self-efficacy towards AI adoption among all other technicians (Shao et al., 2024) and to consider negative impacts on all users, as it will give a holistic picture of other influencing factors during AI adoption.

We want to emphasize the importance of long-term R&D cooperation with SMEs. This is a huge limitation because we cannot safeguard external validity of our results for alternative real-world contexts. Future research directions may include the development of a native or generalizable XAI framework tailored for SMEs, which could provide broader guidance for implementing explainable AI solutions across diverse organizational settings.

ACKNOWLEDGMENTS

The KMI research and development project is part of the funding initiative "Future of Work: Regional Competence Centres for Labour Research – Artificial Intelligence‘ in the programme 'Innovations for Tomorrow's Production, Services and Work" of the Federal Ministry of Research, Technology and Space (BMFTR) for the period from 1 December 2021 to 30 November 2026 and is supervised by the Project Management Agency Karlsruhe (PTKA).

REFERENCES

Akhtar, M.A.K., Kumar, M. and Nayyar, A. (2024). The Role of Human-Centered Design in Developing Explainable AI. In: Towards Ethical and Socially Responsible Explainable AI. Cham: Springer, pp. 99–126. doi:https://doi.org/10.1007/978-3-031-66489-2_4 https://doi.org/10.1007/978-3-031-66489-2_4.

Auernhammer, J. (2020). Human-centered AI: The role of Human-centered Design Research in the development of AI . In: S. Boess, M. Cheung and R. Cain, eds., Synergy - DRS International Conference 2020. pp. 1315–1333. https://doi.org/10.21606/drs.2020.282.

Bunte, A., Richter, F. and Diovisalvi , R. (2021). Why It is Hard to Find AI in SMEs: A Survey from the Practice and How to Promote. In: Proceedings of the 13th International Conference on Agents and Artificial Intelligence. SciTePress, pp. 614–620. https://doi.org/10.5220/0010204106140620.

Campos, J. and Shakhovska, N. (2025). Advancing XAI development: An agile framework for human-centered and explainable AI. In: H. Degen and S. Ntoa, eds. Springer Nature Switzerland, pp. 22–40.

Costa, V.G. and Pedreira, C.E. (2023). Recent advances in decision trees: an updated survey. Artificial Intelligence Review, 56, pp. 4765–4800.

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G. and Ranjan, R. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. ACM Comput. Surv., [online] 55, pp. 1–33. https://doi.org/10.1145/3561048.

Friedrich, J., Brückner, A., Mayan, J., Schumann, S., Kirschenbaum, A. and Zinke-Wehlmann, C. (2024). Human-centered AI development in practice—Insights from a multidisciplinary approach. Zeitschrift Für Arbeitswissenschaft, [online] 78, pp. 359–376. https://doi.org/10.1007/s41449-024-00434-5.

Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., Falco, G., Fiore, S.M., Garibay, I. and Grieman, K. (2023). Six human-centered artificial intelligence grand challenges. International Journal of Human–Computer Interaction, [online] 39, pp. 391–437. Available at: https://doi.org/10.1080/10447318.2022.2153320.

Hein, P., Simon, K., Kögel, A., Löffler, T., Bullinger-Hoffmann, A.C., 2025. Menschzentrierte Einführung von künstlicher Intelligenz in Produktion und Engineering. Zeitschrift für wirtschaftlichen Fabrikbetrieb 120, 12–16. https://doi.org/doi:10.1515/zwf-2024-0147.

Holzinger, A., Carrington, A. and Müller, H. (2020). Measuring the quality of explanations: the system causability scale (SCS) comparing human and machine explanations. KI-Künstliche Intelligenz, [online] 34, pp. 193–198. Available at: https://doi.org/10.1007/s13218-020-00636-z.

Liao, Q.V. and Varshney, K.R. (2021). Human-centered explainable AI (XAI): From algorithms to user experiences. CoRR, [online] abs/2110.10790. Available at: https://arxiv.org/abs/2110.10790.

Merkel-Kiss, M. and Garrel, von (2023). Systematische Literaturanalyse zum KI-Einsatz und KI-basierten Geschäftsmodellen in produzierenden kleinen und mittleren Unternehmen. Zeitschrift für Arbeitswissenschaft, 77, pp. 453–468.

Mo, Y., Zhao, D., Du, J., Syal, M., Aziz, A. and Li, H. (2020). Automated staff assignment for building maintenance using natural language processing. Automation in Construction, [online] 113, p. 103150. https://doi.org/10.1016/j.autcon.2020.103150.

OECD (2025). AI adoption by small and medium-sized enterprises: OECD discussion paper for the G7. [online] Available at: https://doi.org/10.1787/426399c1-en.

Ozmen Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., Falco, G., Fiore, S.M., Garibay, I., Grieman, K., Havens, J.C., Jirotka, M., Kacorri, H., Karwowski, W., Kider, J., Konstan, J., Koon, S., Lopez-Gonzalez, M., Maifeld-Carucci, I., McGregor, S., Salvendy, G., Shneiderman, B., Stephanidis, C., Strobel, C., Ten Holter, C., Xu, W., 2023. Six Human-Centered Artificial Intelligence Grand Challenges. International Journal of Human–Computer Interaction 39, 391–437. https://doi.org/10.1080/10447318.2022.2153320.

Schmager, S., Pappas, I.O. and Vassilakopoulou, P. (2025). Understanding Human-Centred AI: a review of its defining elements and a research agenda. Behaviour & Information Technology, [online] 44, pp. 3771–3810. Available at: https://doi.org/10.1080/0144929X.2024.2448719.

Schwaeke, J., Peters, A., Kanbach, Dominik K, Kraus, S. and Jones, P. (2025). The new normal: The status quo of AI adoption in SMEs. Journal of small business management, [online] 63, pp. 1297–1331. Available at: https://doi.org/10.1080/00472778.2024.2379999.

Shao, C., Nah, S., Makady, H. and McNealy, J. (2025). Understanding user attitudes towards AI-enabled technologies: An integrated model of Self-Efficacy, TAM, and AI Ethics. International Journal of Human–Computer Interaction, [online] 41, pp. 3053–3065. Available at: https://doi.org/10.1080/10447318.2024.2331858.

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction, [online] 36, pp. 495–504. Available at: https://doi.org/10.1080/10447318.2020.1741118.

Yingbo, L., Jianmin, W. and Jiaguang, S. (2007). A machine learning approach to semi-automating workflow staff assignment. [online] Seoul, Korea: Association for Computing Machinery, pp. 340–345. https://doi.org/10.1145/1244002.1244084.