HHAI 2024: Hybrid Human AI Systems for the Social Good
F. Lorig et al. (Eds.)

doi:10.3233/FAIA240210

Investigations to Facilitate Human-Centred AI Development in Manufacturing

Maria Hartikainen
Tampere University, Human Technology Interaction, Unit of Computing Sciences
ORCiD ID: https://orcid.org/0000-0001-5997-0368

Abstract. The integration of artificial intelligence (AI) into various domains offers promising opportunities for innovation and growth. However, alongside these benefits, there is a growing recognition of the potential societal impacts and ethical concerns associated with AI deployment. Addressing these considerations requires a shift towards more human-centred approaches to AI development, emphasising the importance of understanding humans and societal dynamics. My doctoral research explores the concept of Human-Centred Artificial Intelligence (HCAI) especially within the context of smart manufacturing, where AI technologies play a pivotal role in optimising production processes. By facilitating collaboration between humans and AI systems, often referred to as hybrid human-AI intelligence, organisations can enhance productivity while ensuring worker satisfaction and safety. This research proposal outlines a comprehensive investigation into HCAI design and development, aiming to identify novel techniques for integrating human and AI expertise effectively. Through a combination of theoretical analysis and empirical studies conducted in real-world manufacturing environments, this research seeks to contribute to the advancement of HCAI and facilitate the responsible deployment of AI systems in manufacturing and beyond.

Keywords. Human-Centred Artificial Intelligence, Artificial Intelligence, Human-Centred Design, Smart Manufacturing, Human-AI Collaboration, Hybrid Human-AI Intelligence

1. Context

In recent years, there has been a strong integration of artificial intelligence (AI) into various domains and systems [2]. On one hand, AI holds promise of new types of applications and business opportunities [11]. On the other hand, the nature of AI as well as the domains of integration can even lead to undesired societal impact like unfair treatment or discrimination of minority groups or privacy violations [3][11][14]. Addressing these AI-related considerations requires innovative approaches to AI development with more focus on the human factors, and human-AI interaction in diverse application areas [27][30]. Human-Centred Artificial Intelligence (HCAI) is a term referring to various individual, societal, and ethical considerations related to the development of AI [2][5][37]. Following the long traditions of human-centred design (HCD), HCAI emphasises the importance of putting humans and various societal considerations in the centre of the development [2][16]. This includes the basic starting points of understanding user needs and the contextual factors of system design, as well as introduces new ones that are particular to AI as technology [2][8][16][20][24].

Manufacturing is one domain where the integration of AI is strong [6][13]. Demographical reasons and the increasing demand for improved production efficiency are steering the transformation within the manufacturing domain towards smart manufacturing [6][9][26]. In this trend, companies are facing significant technological advancements regarding digitalisation, data analytics, and automation, enabled by advances in AI and the availability of big data [7][14][27][32]. Manufacturing industry is under pressure for more effective production and associated challenge of redesigning both their business processes and their work organisation. Human-AI collaboration is one of the facilitators of more effective productivity in smart manufacturing. In this new form of interaction, humans are paired with AI to work towards common goal [39]. In this collaboration both entities bring their unique strengths and intelligence to enhance the productivity of the collaboration [6][9][10][32]. Therefore, it can be referred as hybrid human-AI intelligence as well [38]. My research aims to contribute to the advancement HCAI and hybrid human-AI intelligence. Specifically, I am investigating what HCAI design and development requires, and novel techniques for integrating human and AI expertise in a way that truly supports effectiveness and enhances worker satisfaction and safety. By combining theoretical analysis with empirical studies in real-world manufacturing environments, I aspire to develop innovative solutions that empower human workers and AI systems to collaborate effectively towards common goals. This research proposal outlines the theoretical framework, research methodology, and potential contributions of the study, followed by a detailed research plan.

The topic of HCAI is relatively novel, however, a large body of literature on topic exists. My work will be based on academic papers related to the main factors of HCAI, like HCAI design [2][3][5][17][20][22][25][27], explainability [1], transparency [18], responsibility [15], fairness [14], ethics [19], and human control [36][37].

In addition to academic papers, the industry [21][24], organisations [11], and designers [8][24] have participated to HCAI design and development by publishing toolkits, guidance, and information on the topic. These tools and recommendations are more practical, aimed for the AI designers and developers. These are going to be acknowledged in my thesis as well.

2. Objectives and research questions

The main objectives of my thesis research are to understand what should be acknowledged in the development of AI that is human-centred. I concentrate on the AI-related factors, investigating what impact they have on the user and the society. Based on this understanding, I propose tools and guidance for the companies in AI design and development, to support them to develop AI solutions that are human centred. In addition, I am set to investigate ways how to support and guide companies in ways that are suitable for different development practises. I am approaching the objectives of this doctoral thesis with two key research questions:

1. What are the considerations, challenges, and relevant building blocks regarding HCAI development?
2. How to facilitate the design of HCAI and human-AI collaborations?

By addressing these research questions, this doctoral study aims to contribute to both academic scholarship and practical integrations of AI for designers and developers. The findings of this research are expected to provide valuable insights for AI designers and developers to facilitate the development of AI solutions that are truly designed for the humans. This can promote the adoption of AI technologies and enhance the effectiveness and trustworthiness of AI systems in the face of AI-enabled future. I will build the answers to the research questions via five studies that support each other and form coherent understanding answering to the main research question. The different studies forming my thesis are described in the next chapter.

3. Studies, methods, and evaluation

I am combining different qualitative research methods to seek answers to my research problem [4][35]. I am using human-centred approach, so I will involve the target users to my studies when possible, and I will try to have at least one iteration rounds for each study. In addition, I am evaluating the proposed tools and guidelines with the target users [16][28].

STUDY 1 – Human-Centred AI Design in Reality: A Study of Developer Companies' Practices
Despite a strong interest in HCAI in academic research, there is little research-based understanding of how the new AI-related requirements and principles manifest in practise in AI development. Hence, my first study has its objectives in understand how HCAI is currently addressed in AI development practises in developer companies. This exploratory research is motivated by the goal of forming new knowledge on how the human-centred design practices are realised in the context of developing AI applications. To gain in-depth understanding of their development practices and to understand how HCAI principles manifest in the current practices of AI development, I conduct an interview study of practitioners from 12 AI developer companies in Finland. Participating companies represent a variety of domain areas and company settings. The AI applications these companies develop vary from automated sensor-based applications to recommendation systems.

STUDY 2 – Towards a Human-Centred Artificial Intelligence Maturity Model
My second study is investigating how to realise HCAI when designing systems that utilise novel computational tools and require consideration of increasingly broad set of requirements, spanning from fairness and transparency to accountability and ethics. I am set to understand what kind of tools could be useful for the companies, so that they could be adoptable and adaptable to various development practises. The objective of this study is to support the HCAI development practices in companies for the humans to have AI solutions that are efficient, trustworthy, and safe. Hence, I propose a maturity model for HCAI (HCAI-MM). In this study, I present the first phase of the model development, in which the central building blocks of HCAI are specified. I then propose model's first phase, in which the central building blocks of HCAI are specified. Initial company requirements for the model's structure and content are evaluated with semi-structured interviews with four AI developers.

Study 3 – Human-AI Collaboration in Smart Manufacturing: Concepts and Framework for Design

With the integration of AI and data in manufacturing, there is a shift from humans merely interacting with technology to actively collaborating with it, especially with AI-enabled agents. This shift brings changes in work practices, tasks, and organisations. I believe that it is important to understand the key concepts to acknowledge and address in the design of effective human-AI collaborations for smart manufacturing. As the objective of this study is to understand and identify what are the key concepts to acknowledge in the design of human-AI collaboration in smart manufacturing, I choose a scoping literature review as a study method. Based on the identified concepts, I propose a human-AI collaboration framework that offers an initial basis for the design of human-AI collaborative systems for smart manufacturing.

Study 4 – Human-AI Collaboration in Smart Manufacturing: Design Method and Guidelines

Human-AI collaboration in smart manufacturing is a growing trend, and this collaboration should be designed appropriately so that it truly effective and supports operator's well-being. In study 4, I am set to understand what kind of design method and design guidelines could support manufacturing companies to design human-AI collaborations. Building on the identified key concepts and the framework of the study 3, I will develop a design method acknowledging the required HCAI factors and the key concepts to address in human-AI collaboration. I will approach the study objectives with interviews with company representatives and designers, in order to understand the user needs. I am evaluating the proposed methods and guidelines with real uses cases from manufacturing companies related to the research topic.

Study 5 – Human-Centred Artificial Intelligence Maturity Model for Smart Industry

HCAI and its requirements might be unfamiliar to manufacturing companies, as they are still finding their ways to work with AI. Main objectives of this study are to investigate how will adoption of AI transform work practises, the organisation, and the society, and to understand how organisations will need to change to accommodate a changing workforce with evolving preferences. By building on the preliminary maturity model from the study 2, I am aiming to build model that is appropriate to practical use and helps companies to reach a higher maturity in HCAI and to support HCAI practises in smart manufacturing. The model provides two perspectives:

1) *"Organisational HCAI Awareness"* – to develop HCAI maturity / HCAI awareness and understanding in the organisation
2) *"HCAI in products and services"* – product development capabilities to develop AI that is human-centred and HCAI ways of working in Smart Manufacturing

Objectives of this study is to support companies to increase their understanding and awareness in HCAI and to support then to reach higher by offering appropriate and concrete guidance how to develop their HCAI capabilities and that way their maturity regarding HCAI. The model will be based on the building blocks presented in the study 2. However, I am adapting that model to smart industry domain, and the domain related factors must be understood and included to the model. In addition, to build a maturity model that truly would be beneficial and useful for the companies, I must understand the company needs. Hence, I will conduct interviews with manufacturing companies and AI developers. The proposed tools I will evaluate with the AI developers and smart

manufacturing companies, to test its usefulness and effectiveness. The evaluation will be conducted via interviews to understand the designer and the company perspective for the proposed model.

4. Preliminary results

The timeline for my thesis January 2022 – December 2025. Figure 1 shows the timeline of my thesis and steps taken and planned steps.

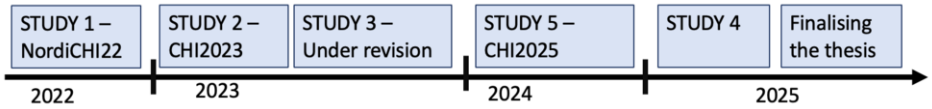

Figure 1. Timeline and the steps of my thesis.

In study 1 I investigated companies' views and practices regarding the development of AI applications, with a specific focus on human-centred viewpoints of the practices. Our study shows that the requirements of HCAI are not well acknowledged currently in AI development. In addition, AI development practices can lean heavily on the experience, methods, and tools familiar from regular SD practices even though data and uncertainty are AI-related factors that require new capabilities, roles, and tasks. These are the real-life struggles we are aiming to create guidance and support for the companies with the following studies. Hence, there is a need for a usable and informatic tool that can be used to support HCAI development in companies.

In study 2, I reviewed related literature and identified the following HCAI factors to acknowledge in HCAI development: (i) *work with AI uncertainty*, (ii) *user control and human-AI collaboration*, (iii) *ethical development and use*: *transparency, accountability, fairness*, and (iv) *trustworthiness: explainability and transparency to build trust between the user and AI*. These factors will serve as a building block of the model created in study 5. In addition, the results of the interviews showed that companies prefer short, compressed information in a check-list form with a possibility to get more information on the topic if needed. They also hope to get targeted tasks and tools related to certain AI product or projects. These insights serve as a starting point for the study 5.

In Study 3, I reviewed related literature and identified six key concepts related to the human-AI collaboration in smart manufacturing: *goal, agent team, skills, task, system,* and *communication*. Based on these concepts, I proposed a human-AI collaboration framework that presents the key concepts and their relations and offers an initial basis for the design of human-AI collaborative systems for smart manufacturing. These concepts serve as a base for developing the human-AI collaboration guidelines in study 4.

5. Discussion

The integration of AI into various domains has been accelerating in recent years, promising new applications and business opportunities while also raising concerns about potential societal impacts such as unfair treatment, discrimination, and privacy violations.

This integration extends deeply into the manufacturing industry, where the pressure for more effective production has catalysed a transformation toward smart manufacturing. Human-AI collaboration emerges as one of the key facilitators in smart manufacturing. This symbiotic relationship, often referred to as hybrid human-AI intelligence, holds the promise of revolutionising the manufacturing landscape by augmenting human capabilities and improving operational efficiency. As companies embrace technological advancements facilitated by AI, there emerges a critical need to ensure that these systems are developed with a human-centric approach, considering the well-being of end-users, and addressing ethical concerns. HCAI offers a framework for addressing these challenges by emphasising the importance of placing humans and societal considerations at the core of AI development. As we explore deeper into the development of HCAI and hybrid human-AI intelligence in the manufacturing sector, several key considerations appear. Firstly, there is a need to develop innovative techniques for integrating human and AI expertise seamlessly. This involves designing AI systems that complement human skills and adapt to diverse working environments, thereby fostering a harmonious collaboration between humans and machines. Additionally, ensuring the effectiveness and safety of human-AI interactions requires comprehensive training and support mechanisms for workers to navigate these novel work arrangements successfully. My research aims to address these challenges and contribute to the advancement of HCAI and hybrid human-AI intelligence in manufacturing. By combining theoretical insights with empirical studies conducted in real-world manufacturing environments, I seek to develop actionable solutions that empower human workers and AI systems alike. Through a holistic approach encompassing HCD, ethical considerations, and technological innovation, I aim to pave the way for a future where humans and AI collaborate synergistically, driving sustainable growth and innovation in the manufacturing industry. By prioritising workers well-being, ethical considerations, and effective human-AI collaboration, we can harness the transformative potential of AI while mitigating its associated risks, thereby creating a more inclusive, sustainable, and productive future for all stakeholders involved.

5.1. Future work

Building on the insights of the studies included to my doctoral research, in the future it would be beneficial to develop comprehensive toolkit for the HCAI design in manufacturing domain. This toolkit should include instructions, tools, and concrete guidelines to support appropriate HCAI and human-AI collaboration to fully benefit the new technological systems and to offer positive user and customer experience. Objective is to map and understand HCD requirements in manufacturing and to understand how to support companies to understand and gather their clients' needs for a more tailored, engaging experience for their customers.

Acknowledgement

Supervisors: Kaisa Väänänen and Thomas Olsson, Tampere University
Funding: Study 1 and study 2 are funded from European Regional Development Fund, Cohesion Fund and European Social Fund (fund number 306672). Study 3, study 4, and study 5 are funded from Business Finland -funding (grant 2451/31/2022)

References

[1] Arrieta, Alejandro Barredo, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García et al. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI. Information Fusion 58 (2020): 82-115. DOI:10.1016/j.inffus.2019.12.012

[2] Auernhammer, Jan. 2020. Human-centered AI: The role of Human-centered Design Research in the development of AI. In Synergy - DRS International Conference 2020. Held online. DOI:10.21606/drs.2020.282

[3] van Berkel, Niels, Tag, Benjamin, Goncalves, Jorge, and Hosio, Simo. 2020. Humancentred artificial intelligence: a contextual morality perspective. Behaviour & Information Technology. 1-17.

[4] Braun, Virginia, and Clarke, Victoria. 2012. Thematic analysis. APA handbook of research methods in psychology, Vol. 2: Research designs: Quantitative, qualitative, neuropsychological, and biological. American Psychological Association, Washington. DOI: 10.1037/13620-000

[5] Bond, Raymond R., Mulvenna, Maurice D., Wan, Hui, Finlay, Dewar D., Wong, Alexander, Ansgar, Koene, Brisk, Rob, Boger, Jennifer and Adel, Tameem. 2019. Human Centered Artificial Intelligence: Weaving UX into Algorithmic Decision Making. In Proceedings of 16th International Conference on Human-Computer Interaction. Bucharest, 2-9.

[6] Breque, M., De Nul, L., & Petridis, A. (2021). Industry 5.0: towards a sustainable, human-centric and resilient European industry. Luxembourg, LU: European Commission, Directorate-General for Research and Innovation.

[7] Cantamessa, Marco, Montagna, Francesca, Altavilla, Stefania, and Casagrande Seretti, Alessandro. 2020. Data-Driven Design: The New Challenges of Digitalization on Product Design and Development. Design Science, 6. DOI: 10.1017/dsj.2020.25

[8] Cerejo, Joana. 2021. The design process of human-centered AI. Retrieved Feb 8, 2022, from https://bootcamp.uxdesign.cc/human-centered-ai-design-processpart-1-8cf7e3ce00

[9] Cortés-Leal, A., Cárdenas, C., & Del-Valle-Soto, C. (2022). Maintenance 5.0: towards a worker-in-the-loop framework for resilient smart manufacturing. Applied Sciences, 12(22), 11330.

[10] Dellermann, Dominik, Adrian Calma, Nikolaus Lipusch, Thorsten Weber, Sascha Weigel, and Philipp Ebel. 2021. The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems. arXiv preprint arXiv:2105. 03354.

[11] Dignum, Virginia. 2019. HumanE AI. Project report. Retrieved Oct 8, 2021, from https://www.humane-ai.eu/wp-content/uploads/2019/11/D13-HumaneAIframework-report.pdf [16] Dove, Graham, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX design innovation: Challenges for working with machine learning as a design material. In Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, Neew York, 278-288. DOI: 10.1145/3025453.3025739

[12] Eiband, M., Völkel, S. T., Buschek, D., Cook, S., & Hussmann, H. (2019, March). When people and algorithms meet: User-reported problems in intelligent everyday applications. In Proceedings of the 24th international conference on intelligent user interfaces (pp. 96-106).

[13] El Hamdi, S., Abouabdellah, A., & Oudani, M. (2019, June). Industry 4.0: Fundamentals and main challenges. In 2019 International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA) (pp. 1-5). IEEE.

[14] Feuerriegel, Stefan, Mateusz Dolata, and Gerhard Schwabe. 2020. Fair AI: Challenges and Opportunities. Business & information systems engineering 62, 4 (2020): 379-384. DOI: 10.1007/s12599-020-00650-3

[15] Ghallab, Malik. 2019. Responsible AI: requirements and challenges. AI Perspectives, 1:1. 1-7. DOI: 10.1186/s42467-019-0003-z

[16] Giacomin, Joseph. 2014. What is human centred design? The Design Journal 17, 4 (2014): 606-623. DOI:10.2752/175630614X14056185480186

[17] Girardin, Fabien, and Neal Lathia. 2017. When user experience designers partner with data scientists. In AAAI Spring Symposium Series on Designing the User Experience of Machine Learning Systems. The IEEE press, Palo Alto.

[18] Guidotti, Riccardo, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. 2018. A survey of methods for explaining black box models. ACM computing surveys 51, 5 (2018). 1-42.

[19] Hagendorff, Thilo. 2020. The ethics of AI ethics: An evaluation of guidelines. Minds and Machines 30, 1 (2020). 99-120. DOI:10.1007/s11023-020-09517-8

[20] Harper, Richard. 2019. The Role of HCI in the Age of AI. International Journal of Human–Computer Interaction, 35, 15 (2019). 1331-1344. DOI: 10.1080/10447318.2019.1631527

[21] HAX. 2021. Microsoft Human-centered AI eXperience Toolkit. Retrieved Feb 8, 2022, from https://www.microsoft.com/en-us/haxtoolkit/

[22] Hill, Charles, Bellamy, Rachel, Erickson, Thomas, and Burnett, Margaret. 2016. Trials and Tribulations of Developers of Intelligent Systems: A Field Study. In 2016 IEEE Symposium on Visual Languages and Human-Centric Computing.162-170. DOI: 10.1109/VLHCC.2016.7739680
[23] Holmquist, Lars Erik. 2017. Intelligence on tap: artificial intelligence as a new design material. Interactions 24, no. 4 (2017). 28-33. DOI:10.1145/3085571
[24] IBM Design for AI. 2019. Retrieved Oct 8, 2021 from https://www.ibm.com/design/ ai/
[25] Jin, Xiaoneng, Evans, Mark, Dong, Hua, and Yao, Anqi. 2021. Design Heuristics for Artificial Intelligence: Inspirational Design Stimuli for Supporting UX Designers in Generating AI-powered Ideas. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. 1-8. DOI: 10.1145/3411763.3451727
[26] Jwo, J. S., Lin, C. S., & Lee, C. H. (2021). Smart technology–driven aspects for human-in-the-loop smart manufacturing. The International Journal of Advanced Manufacturing Technology, 114, 1741-1752.
[27] Lu, Jiahao, Ortega, Alejandra. G., Gonçalves, Milene, and Bourgeois, Jacky. 2021. The impact of data on the role of designers and their process. Proceedings of the Design Society, 1. 3021-3030. DOI: 10.1017/pds.2021.563
[28] Maguire, Martin. 2001. Methods to support human-centred design. International journal of human-computer studies 55, 4 (2001). 587-634. DOI:10.1006/ijhc.2001.0503
[29] Mittal, S., Khan, M. A., Romero, D., & Wuest, T. (2019). Smart manufacturing: Characteristics, technologies and enabling factors. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 233(5), 1342-1361.
[30] Oppermann, Leif, Alexander Boden, Britta Hofmann, Wolfgang Prinz, and Stefan Decker. 2019. Beyond HCI and CSCW: Challenges and useful practices towards a human-centred vision of AI and IA. In Proceedings of the Halfway to the Future Symposium. ACM, 1-5. DOI: 10.1145/3363384.3363481 [36] PAIR with Google. People + AI Research. Retrieved Oct 8 from https://pair. withgoogle.com.
[31] Pappas, I. O., Mikalef, P., Dwivedi, Y. K., Jaccheri, L., & Krogstie, J. (2023). Responsible Digital Transformation for a Sustainable Society. Information Systems Frontiers, 1-9.
[32] Passalacqua, M., Pellerin, R., Doyon-Poulin, P., Del-Aguila, L., Boasen, J., & Léger, P. M. (2022, June). Human-Centred AI in the Age of Industry 5.0: A Systematic Review Protocol. In International Conference on Human-Computer Interaction (pp. 483-492). Cham: Springer Nature Switzerland.
[33] Passi, Samir and Jackson, Steven J. 2018. Trust in Data Science: Collaboration, Translation, and Accountability in Corporate Data Science Projects. Proceedings of the ACM on Human-Computer Interaction, 2. 1-28. DOI: 10.1145/3274405
[34] Piorkowski, David, Park, Soya, Wang, April Y., Wang, Dakuo, Muller, Michael, and Portnoy, Felix. 2021. How AI Developers Overcome Communication Challenges in a Multidisciplinary Team: A Case Study. Proceedings of the ACM on HumanComputer Interaction, 5. 1-25. DOI: 10.1145/3449205
[35] Saldaña, Johnny. 2014. Coding and analysis strategies. In The Oxford handbook of qualitative research. DOI: 10.1093/oxfordhb/9780199811755.013.001
[36] Shneiderman, Ben. 2020. Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy Human-Centered AI systems. ACM Transactions on Interactive Intelligent Systems, 10, 4 (2020). 1-31. DOI: 10.1145/3419764
[37] Shneiderman, Ben. 2020. Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. International Journal of Human–Computer Interaction 36, 6 (2020). 495-504. DOI: 10.1080/10447318.2020.1741118
[38] Schleiger, E., Mason, C., Naughtin, C., Reeson, A., & Paris, C. (2023). Collaborative Intelligence: A scoping review of current applications. Qeios.
[39] Xu, W., Dainoff, M. J., Ge, L., & Gao, Z. (2023). Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI. International Journal of Human–Computer Interaction, 39(3), 494-518.