

Reviewing Human-Centric Aspects of AI Collaboration in Manufacturing

Downloaded from: https://research.chalmers.se, 2026-11-08 03:51 UTC

Citation for the original published paper (version of record):
Jaksic, S., Notarianni, A., Salunkhe, O. et al (2027). Reviewing Human-Centric Aspects of AI Collaboration in Manufacturing. IFIP Advances in Information and Communication Technology, 812 IFIPAICT: 140-154. http://dx.doi.org/10.1007/978-3-032-38610-6_10

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Reviewing Human-Centric Aspects of AI Collaboration in Manufacturing

Sandra Jakšić[1(✉)], Anita Clara Notarianni[1], Omkar Salunkhe[1], Ana Correia Simões[2], Anna Syberfeldt[3], and Johan Stahre[1]

[1] Chalmers University of Technology, Gothenburg, Sweden
sandra.jaksic@chalmers.se
[2] INESC TEC, Porto, Portugal
[3] University of Skövde, Skövde, Sweden

Abstract. The increasing integration of artificial intelligence (AI) in manufacturing is reshaping human roles on the shop floor. While AI has the potential to enhance human capabilities, questions arise about its impact on workers. Ensuring human-centricity remains crucial, but there is a need to investigate how these principles translate to human-AI collaboration. This paper presents a structured literature review of 25 papers, examining human-centric goals, challenges and enablers in human-AI collaboration. The results show that the most emphasized goals include trust, transparency, augmentation, explainability, safety, and well-being. Prominent challenges include job displacement, conflicting priorities, and algorithmic bias. Identified enablers include co-design approaches, explainable interfaces, and dynamic task allocation. The findings highlight that human-centric aspects are interconnected, requiring a holistic approach. Although human-centricity is widely advocated, it is often framed at the system interface level, rather than in organizational or design processes. Future research should deepen the understanding of cognitive and psychosocial dimensions of human-AI collaboration and develop more empirical and longitudinal insights.

Keywords: Human-centricity · Human-AI collaboration · Manufacturing · Industry 5.0 · Human Factors

1 Introduction

The rapid growth in artificial intelligence (AI) is reshaping manufacturing industry by enabling a new form of adaptive decision support and flexible automation [1]. AI is a multidisciplinary field focused on creating systems and algorithms, capable of performing tasks and actions that typically require human intelligence [2]. In many visions of industry 4.0 and industry 5.0, human-AI collaboration is seen as the path forward to resilient, sustainable and flexible systems [3]. As use of AI systems increases on the shop floor, critical questions arise on how AI can support human needs, as humans are still the most flexible resource on shop floor. Human-centricity therefore emerges as a central aspect in this context.

© The Author(s) 2027
Published by Springer Nature Switzerland AG 2027
G. Pezzotta et al. (Eds.): APMS 2026, IFIP AICT 812, pp. 140–154, 2027.
https://doi.org/10.1007/978-3-032-38610-6_10

Human-centric approaches emphasize that technological systems should be designed around human values, capabilities and limitations [4]. Beyond ergonomics and safety, such approaches also promote inclusivity, empowerment and opportunities for self-development and growth [5]. Furthermore, centering the human in technology and work system design promotes acceptance of technology, which is particularly crucial as human resistance and mistrust towards new technologies have been identified as one of the main barriers in digital transformation [6]. Human-centred AI (HCAI) extends these principles by focusing on AI systems that augment human intelligence with machine intelligence [7, 8]. This paradigm emphasizes human well-being and alignment with human values, ethics, and needs [9].

Many human-centric themes in AI build on earlier human-automation literature. One study, [10], introduced a taxonomy of automation levels and functions, and identified human performance consequences in terms of mental workload, complacency and skill degradation that result from function allocation. Another study, [11], defined situation awareness and showed how automation can erode operators' ability to perceive, comprehend, project system states, and assume an active role in the system. A 2020 review of human-autonomy teaming [12] further shows how the themes of agent transparency, reliability, task characteristics, automation level, team composition and training have been investigated in the past decades on the way they shape collaboration outcomes. However, it is unclear how human-centric principles will translate into the new domain of human-AI collaboration in manufacturing, especially since AI is a dynamic and increasingly autonomous, but yet opaque, partner. Different concepts have been floated, such as "human-in-the-loop" [13] and "hybrid intelligence" [14], that cover human-AI collaboration, but such concepts primarily emphasize productivity and performance gains.

The theme of HCAI in manufacturing has recently attracted several reviews. Prior reviews have examined psychosocial and cognitive human factors in Industry 4.0/5.0 contexts [15], conceptual clarifications [16], and technological enablers across the product lifecycle [17]. However, these reviews rarely explicitly structure findings around the interplay between human-centric goals, challenges, and enablers that allow industry to form actionable strategies for change. Therefore, the aim of this review is to examine recent literature on human-AI collaboration in manufacturing to identify contemporary human-centric goals, the challenges that hinder them, and the enablers that support them. Based on this aim, the research question formulated for the study is: What human-centric goals, challenges, and enablers should be considered in the context of human-AI collaboration in manufacturing?

2 Method

This study employs a structured literature review to examine human-AI collaboration in manufacturing. The search was conducted in Scopus, one of the largest multidisciplinary commercial databases, using search terms related to 1) human-centricity, 2) human-AI collaboration, and 3) manufacturing. The following search string was used:

TITLE-ABS-KEY (("human-centric" OR "human-centered" OR "human-centred") AND ("human-AI collaboration" OR "human-AI teaming" OR "human-AI symbiosis"

OR "human-AI cooperation") AND (manufacturing OR "assembly" OR "shop floor" OR "production" OR "industry 5.0")) AND PUBYEAR >2020 AND PUBYEAR <2027 AND (LIMIT-TO (LANGUAGE,"English")) AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"cp") OR LIMIT-TO (DOCTYPE,"re"))

As the AI field moves forward in a breathtaking speed, we delimited our study to a narrow time window, 2021–2026. The selected timeframe reflects the emergence of the Industry 5.0 paradigm and its increased focus on human-centricity [18], while also capturing recent AI developments. This allowed us to focus on current AI applications rather than exploring the development of the field.

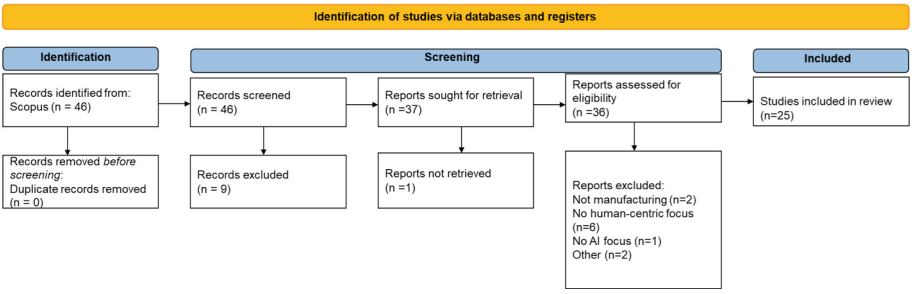


Fig. 1. Structured literature review process

The query retrieved 46 records. Following title and abstract screening based on predefined inclusion criteria, 37 papers were selected for full-text review. The following inclusion criteria were used: 1) addressing human involvement, behaviour or activities, 2) analyze human interaction with AI systems, and 3) in a manufacturing context. Studies focusing solely on technological development were excluded. After full text screening, 25 papers were included in this review. The screening process is shown in Fig. 1

The corpus was analyzed through thematic analysis. Each paper was independently coded by two authors, and disagreements were discussed until consensus was reached. Record were coded for AI type, domain application, AI task, human role, as well as human-centric goals, challenges, and enablers. These high-level categories were predetermined to align with the research question, and used as an analytical lens rather than a mutually exclusive taxonomy. Goals refer to desired human-centric outcomes of human–AI collaboration, challenges represent barriers or risks that may hinder the achievement of these outcomes, and enablers denote mechanisms, system features, or organizational practices that support the goals or mitigate challenges. While some concepts appear across categories, they are interpreted differently depending on their role within the system.

2.1 Application Context

Types of Tasks. The corpus spans several manufacturing application domains, reflecting AI's broad potential to support diverse functions and tasks across the industry. The most common application domain is quality control and inspection. In this sector, AI is

deployed to enhance inspection, defect detection and quality prediction processes [19–22]. Within quality control settings, AI-based operator training is also explored [23]. Another commonly recurring application is assembly [24, 25], where AI applications are often presented in combination with collaborative robotics systems, covering scenarios ranging from human–robot collaborative task scheduling and cobot systems that dynamically provide operators with the correct tools during manual assembly.

Maintenance applications are also described, particularly the predictive maintenance domains [26, 27]. AI systems are used to anticipate equipment failures before they occur and estimate residual useful life to reduce unplanned downtimes.

AI systems are also used in production management contexts, as decision support for resource allocation and production optimization [28, 29].

From the corpus, more specialized manufacturing contexts also emerge. For instance, Large Language Models (LLMs) are used in warehousing [17, 30]. Furthermore, [26] explores generative design algorithms used to optimize products for disassembly and recycling, and [31] investigates the use of AI-driven approaches for hazard identification in Industry 5.0.

Types of AI. Deep learning models are the most used AI types across the papers. They are applied to perception tasks such as visual inspection and defect detection [20–22], as well as predictive tasks [27]. As these models have a black-box nature, they are often paired with explainability and interpretability methods to enhance transparency and trust [20–22, 29, 32, 33]. LLMs are also a recurring technology in the corpus, as their capabilities can be leveraged to enable natural language interactions [30], and are thus ideal for tasks such as report generation, natural language explanation of inspection data [19] and to serve as digital assistants. Some papers employ or mention hybrid approaches that combine data-driven learning with symbolic reasoning, such as Neuro-Symbolic Learning [33]. In fields such as assembly, where robotics applications have long been implemented, applications of embodied AI are reported [24], where AI is integrated into physical entities such as robots, giving them the ability to interact and perceive the physical surroundings. Reinforcement learning algorithms are also applied to planning and scheduling problems [26].

3 Results

The results are structured around three categories: Goals, Challenges, and Enablers. A brief overview of the categories and the most frequent human-centric aspects are given in Table 1. The reported frequencies refer to how many of the included papers mention each aspect.

3.1 Goals

Goals are crucial for motivating industrial guidelines when designing human-centred systems. Several studies mentioned augmentation of human capabilities rather than human replacement as an objective in AI implementation [17, 19, 26, 32, 36]. Similarly, studies reported the need for complementarity between human and AI capabilities [29]. By

Table 1. Most frequent goals, challenges, and enablers identified in the review

Main aspects	References	Freq.
Goals		
Trust	[17, 21, 24, 27, 32, 34]	6
Transparency	[17, 21, 24, 27, 34, 35]	6
Augmentation	[17, 19, 26, 32, 36]	5
Human-AI capability complementarity	[26, 29, 37, 38]	4
Explainability	[17, 20, 34, 35]	4
Safety	[26, 31, 32, 36]	4
Wellbeing	[17, 26, 31, 36]	4
Challenges		
Job displacement	[17, 31, 36, 39]	4
Conflicting priorities	[17, 31, 32, 36]	4
Algorithmic bias	[17, 26, 28]	3
Skill gaps and lack of training	[23, 36, 40]	3
Enablers		
Human oversight	[19, 25, 26, 30, 32, 35]	6
Co-design and stakeholder inclusion	[31–33, 41, 42]	5
Explainable interfaces	[19–21, 24, 25]	5
Dynamic task allocation	[25, 28, 37]	3

combining the flexibility of human intelligence with the scalable processing power of AI, researchers argued for the development of more flexible systems [37]. On one hand, human contextual knowledge, intuition and feedback can be incorporated to enhance the system performance in terms of efficiency, accuracy, and adaptability, as stated by [26]. On the other hand, [41] argued that human empowerment arises from providing humans with access to up-to-date tacit knowledge, while AI systems leverage this same knowledge base to generate new, trustworthy insights in support of human decision-making. One study also suggested that technical systems may improve accessibility to a previously underqualified workforce [41].

Collaboration was also presented as a goal to leverage both human and non-human capabilities to improve operational performance [38]. Studies reported the need for considering the interaction between humans and AI [17, 28, 33], effective integration of actors [28], and communication [38] as elements in human-AI collaboration. Safety is also linked to this, as effective human-AI interaction was described as necessary for ensuring safe operation [31].

Skills were also identified as a factor in human-AI collaboration by [38]. Another study [23], evaluated motivation and engagement for developing longevity and success of skill acquisition, to be able to react in cases of AI failures.

Multiple studies emphasized well-being as a priority in human-AI collaboration [17, 26, 31, 36]. This includes safety, health and satisfaction, according to [36]. Safety was also reported as a factor by [26, 31, 32]. User needs and interaction between agents are to be considered to identify potential risks or impacts [33]. Safety entails considering both physical and psychological risks, as described by [32]. Beyond safety, the alignment with human values, needs, preferences and limitations was also mentioned by [17], where the AI understands human behavior, values and emotions. One study [24] considered minimizing the cognitive load, while another [25] mentioned increasing acceptance by minimizing frustrations through responsive systems. Operator satisfaction can serve as an evaluation metric for human-AI collaborative frameworks [35]. Additional approaches promoted work-life balance through remote work and on-demand learning [41]. Notably, [34] found that stakeholders have different priorities regarding human-AI collaboration, where employers value the potential productivity gains, while employees highlight seamless cognitive collaboration and improved working conditions. The consideration of human cognitive capabilities was also mentioned by [17].

Ethical considerations were reported in the context of human-AI collaboration systems [17, 32, 35]. According to [17], ethical considerations include privacy protection and fairness, while [35] proposed that transparency, explainability, and fairness should be addressed for reducing ethical concerns. One study [32] proposed an ethical framework for creating trust, including the following principles: human agency, transparency, privacy, fairness, safety, accountability, and user involvement.

Trust and transparency emerged as requirements in several studies [17, 21, 24, 27]. In addition, explainability [17, 20], reliability [17], interpretability [17], and understandability [17]. According to [34], explainability, transparency, and reliability are trust-building factors. One study, [24], further highlighted trust building across multiple stakeholder groups, including operators, developers, and certifiers.

3.2 Challenges

Although Challenges Encountered and Documented in the Reviewed Papers May Serve as Excellent Drivers for Change, it Appears that Prevalent Challenges Mentioned Are Primarily towards the Threat Side of the Spectrum and Should Therefore Be Seen as Such. Job Displacement Was Identified as a Concern by [17, 31, 36, 39]. Employees in Developed Countries Were Found to Be Broadly Familiar with AI [39], Attributed to the Prevalence of Cognitive Work, but a Lack of Complementarity Between Human Skills and AI Capabilities Could Lead to Job Displacements. Developing Countries Are less Subject to AI-Driven Displacement Due to a Higher Proportion of Physical Jobs, which Hinders AI Adoption [39]. Digital Inequality Amplification Was Also Reported by [17, 36].

Lack of worker training in parallel with AI implementation was found to be a hinderance to successful collaboration [40]. The changing skill needs in high automation tasks was acknowledged by [23], where humans need to be able to intervene in cases on AI failure. A digital skill gap was described by [36], where the need for basic cognitive and physical skills might be reduced, but the ability to think critically, digital literacy, and the ability to work with AI will be needed. Fully automated AI was shown by [23] to negatively impact skill acquisition, perceived autonomy, motivation and engagement.

Several studies mentioned the challenge of conflicting priorities [17, 31, 32, 36]. This was described by [17], as a challenge for value alignment between humans and AI. Another study [31] gave an example of system properties conflicting with safety, and trade-offs between security and safety. Other ethical conflicts were listed by [32], namely, navigating transparency but at the same time not exposing personal data, ensuring human agency but without compromising safety, ensuring accountability but respecting privacy, inclusivity through individual accommodations but not reducing predictable behavior and safety, or giving autonomy but not introducing human bias.

Algorithmic bias was mentioned as a challenge by [17, 26, 28]. The challenge of lack of human oversight was mentioned by [27], and similarly human-AI control and responsibility by [17]. Additionally, one study [36] reported rebound effects as a risk.

Several AI safety challenges were listed by [31], including unclear hazard identification, the limits of purely technical analysis, system properties conflicting with safety, the need for transdisciplinary expertise, concerns about human displacement, trade-offs between security and safety, risks arising from human errors and sensor faults. The same study also reported that new technologies such as AI are introducing new types of challenges, leading to difficulties in risk identification, for example, due to reduced standardization due to the nature of adaptable systems. Changing workforce also introduces new types of challenges, as mentioned by [31], such as the inclusion of vulnerable groups, e.g. people with disabilities, the elderly, and migrant workers.

3.3 Enablers

Finally, the reviewed papers proposed several enablers for addressing the identified goals and challenges. Dynamic task allocation between human and non-human actors was frequently addressed in the literature to improve operational performance. Studies considered skills and roles [37], as well as cognitive and physical task complexity [28]. For example, [25], integrated dynamic scheduling into a digital twin platform where AI-enhanced robots adapt task execution in real time to reduce physically demanding and repetitive work for operators.

The need for continuous upskilling and reskilling of the existing workforce, in addition to hiring new talent, was noted by [36]. A virtual assistant was used by [28] to enhance worker capabilities, and highlights the importance of streamlining onboarding and continuously enhancing human capabilities to improve the interaction with the system. The virtual assistant provides customized guidance to the operator based on skill level and can also support operators in handling equipment failures by providing troubleshooting recommendations based on machine data and documentation from past operator experiences. It retrieves human-generated, domain-relevant documentation to mitigate LLM biases. It is further complemented by a digital twin for virtual solution testing, improving human understanding of the impacts of changes before implementation.

Studies demonstrated improved explainability through visual and textual communication. Proposed approaches include operator dashboards [21], visual displays [25], saliency maps [20], spider charts [24], natural language explanations [19, 21], and traceability logs [32]. For example, [21] described an operator dashboard supporting explanations, manual overrides, feedback injection, and traceability of actions, while [25]

enabled workers to understand and intervene in robot actions through visual communication of robot intentions. In contrast to images, text has also been used for explainability [19, 21]. Descriptions in text form were used by [19] as a complement to enhance the interpretation and validation of quality analysis results. Similarly, [21] integrated an interpretable decision tree model, providing human-readable root cause explanations. Natural language summaries are created based on decision-tree rules and contextual information used by the AI model. However, it was found that workers with preexisting expertise demonstrated higher performance outcomes when supported by explainable AI, compared to workers with lower experience [20]. Explainability enables experts to utilize their domain knowledge, to determine when to rely on the system and when to override, leading to improved operational performance [22]. One study [32] suggested creating auditability by keeping logs for tracing back actions, recommendations, and overrides.

Many of these explainability mechanisms were also mentioned as trust-enabling factors. Explainable AI, transparency, human-in-the-loop, semantic agents, and decision support were suggested by [27], to address challenges related to trust and human oversight. Another study [24] aimed at building trust for embodied AI agents by ensuring explainability, predictability, and environmental awareness. One study, [32], further described trust as a lifecycle consideration embedded throughout AI design, validation, deployment, operation and evaluation. The same study highlighted transparency of system actions, privacy-preserving data practices, and integrated worker training as mechanisms for sustaining trust.

Multiple studies highlighted human oversight in AI collaboration. Such systems can provide workers with the option to intervene or override AI plans [25], particularly in complex decisions requiring human contextual judgement [35], or in unexpected events in which human-in-the-loop approaches enable the integration of qualitative considerations [26]. Additional mechanisms included feedback loops enabling workers to give input for future iterations [19], and worker-guided AI optimization to incorporate worker priorities [30]. Human oversight was further associated with accountability [32], fairness against algorithmic bias [26], with research, [23], also demonstrating that AI decision aids with final human authority enhance worker motivation, engagement and skill acquisition.

One study [32] suggested addressing fairness by embedding bias-mitigation techniques to reduce systemic disadvantages, and addressing inclusivity by ensuring accessibility and effective performance from diverse workers, such as those with varying abilities or skill levels. In addition, the evaluation of ethical performance encompassed metrics such as user trust levels, usage and override statistics, incidents and near misses, complaints and feedback tickets, diversity and fairness metrics, adoption and retention, and performance with ethical constraints [32].

One study [17] placed humans at the center throughout design, development and use of AI systems. More broadly, stakeholder inclusion emerged as a recurring aspect in the development of human-AI collaboration systems [31–33, 41, 42]. Stakeholder inclusion in impact analysis and in choosing safeguards was proposed by [32] as a way of addressing conflicts between ethical principles, as different stakeholders might have different perspectives on ethical challenges, with the aim of also identifying concerns regarding

trust and ethics. The need for including diverse stakeholder perspectives in risk identification and addressing safety challenges was stated by [31, 33] as well. User-centered approaches were used to integrate worker perspectives through methods such as interviews, workshops, observations, and prototype testing [33, 42]. These approaches were reported to support knowledge sharing and the development of shared understanding of the desired solution [33].

4 Discussion

The review highlights a recurring set of human-centric goals in human-AI collaboration in manufacturing, as shown in Table 1. Trust [17] and transparency [21] are most frequently mentioned, followed by augmentation [19], capability complementarity [26], explainability [20], safety [31], and well-being [36]. Collectively, these goals suggest that human-centric AI is expected not only to improve operational performance, but also to shape human experience, understanding and preserve human control of AI-assisted work.

Transparency, explainability and safety can influence trust, and at the same time, trust, transparency and explainability are often treated as components of a broader ethical consideration, positioning ethics as a meta-goal. Safety similarly operates both as a direct objective and as a component of well-being, while also shaping trust. This interconnect structure implies that isolated human-centric objectives are insufficient. Consequently, our findings indicate that human-centric aspects are interconnected rather than independent. Also indicated, is that effective human-centric AI requires holistic design approaches that explicitly and carefully manage the interdependencies and trade-offs mentioned above.

Job displacement emerges as a frequently mentioned challenge [39], potentially linked to emerging skill gaps and a lack of capability complementarity between AI and human skills. Not surprisingly, dynamic task allocation, considering human aspects and tasks, appeared as a recurring enabler [37]. The findings support capabilities-based division where AI handles data-intensive and repetitive tasks, while humans contribute with contextual judgement, creativity, social interaction, and managing unexpected events. Task division mechanisms should therefore be embedded in the design of human-AI systems. However, the review also suggests that "balancing" human and AI contributions is not sufficient on its own. Several studies emphasize training, onboarding, and continuous learning as necessary conditions for sustained and safe collaboration. At the same time, the literature describes skill requirements and organizational changes only loosely, indicating a gap between proposed human-centric principles and actionable implementation guidance.

Moreover, literature projection of "human-centricity" appears to be skewed towards physical ergonomics and reducing immediate physical risks. While AI-enabled robotics and automated solutions can effectively reduce physical strain on human operators, the cognitive and psychosocial dimensions of a particular task or work receive comparatively less attention. Specifically, the findings indicate that current applications tend to address short-term physical adaptation more than the long-term effects on human cognition, such as cognitive load, skill degradation, overreliance, and changes to decision-making

practices. These gaps point to a need for deeper investigation of cognitive ergonomics and psychosocial impacts in human-AI collaboration.

Conflicting priorities were also identified as a challenge which affect ethics [32]. These issues are compounded by challenges at the organizational levels. Co-design of human-AI collaboration systems appears as a highly mentioned enabler [33]. The method is used to address conflicting priorities [32] and ensure safety [31]. An important finding of this review is how human-centricity is interpreted by different stakeholders. For some, human-centricity is protecting human operators from physical harm and ensuring their safety. For others, human-centricity emphasizes inclusion, empowerment and meaningful work. It is important to recognize these different aspects through all phases from development to implementation.

Algorithmic bias emerges as another prevalent challenges [17], while explainable interfaces appear as a common enabler for addressing explainability and transparency [20]. The findings suggest that human-AI collaboration is not only a technical integration challenge but also a socio-cognitive one, where users must understand and rely on system outputs. While prior Industry 5.0 research highlights similar priorities, the present review indicates that such aspects may be predominantly addressed at interface level. A possible explanation for this is that interface-level mechanisms, such as explainability and transparency are easier to operationalize and evaluate on engineering-level research than on organizational-level processes that include governance or skills development. The organizational conditions that sustain trust over time, such as governance, continuous training, and worker involvement, receive comparatively little attention. This may indicate a gap between system design and organizational implementation, pointing to the need for integrated socio-technical approaches.

The findings in this study are consistent with previous literature on human-centric AI and Industry 5.0 in terms of the emphasis on trust, transparency, and human oversight. Previous reviews have focused on conceptual definitions, cognitive and psychosocial factors, or technological enablers. In contrast, this study specifically structures the literature around human-centric goals, challenges, and enablers in human-AI collaboration in manufacturing. By doing so, this review contributes towards guidance in designing AI-enabled manufacturing systems by clarifying desired outcomes and risks, as well as providing examples of enabling mechanisms, to create systems that are not only efficient but also aligned with human needs and values.

This study provides relevant implications for both research and practice. From a research perspective, it indicates the need to move beyond conceptual discussions of human-centricity towards empirically validated frameworks, measurable indicators, and longitudinal analyses of human-AI collaboration. From a practical perspective, the results suggest that organizations should adopt holistic implementation approaches that consider not only technical performance but also human factors such as trust, skill development, cognitive load, and well-being. Integrating explainability, human oversight, and continuous training into system design may support more effective and sustainable human-AI collaboration.

4.1 Identified Limitations

The use of a single database (Scopus) may introduce selection bias, as relevant studies indexed in elsewhere could have been excluded. However, Scopus coverage has historically a well-documented and higher coverage of manufacturing-related research. The specificity of the search string, while ensuring focus, may have limited the inclusion of studies addressing human-AI collaboration under alternative terminologies (e.g., human–machine interaction or computer vision). We decided to avoid sub-groups of AI to capture the contemporary discussion on AI applications in manufacturing, this was in line with the narrow time-span of the review. Further work should naturally address adjacent and complementary areas. When this limitation addressed through a sensitivity check with additional, broader search terms, we found more than 350 papers, with many of these addressing deeper technological development or primary objectives in different industrial domains, e.g. Air Traffic Control. The limited corpus ($n = 25$) may constrain the generalizability of our findings and may capture only parts of the diversity of applications and perspectives across the broad manufacturing domain, but it does provide an exploratory overview of emerging themes and priorities in recent literature.

Although the review coding process was conducted by two of the authors, no formal inter-coder reliability metric was reported, which would introduce subjectivity in the thematic classification. Finally, as this study is based on a literature review, the relationships identified between human-centric goals, challenges, and enablers are not empirically validated, and should therefore be interpreted as conceptual rather than causal.

4.2 Future Research

Future research should focus on operationalizing human-centricity by developing measurable indicators and evaluation frameworks for constructs such as trust, transparency, well-being, and human autonomy in AI-assisted work environments. Further investigations are also needed at the organizational level to understand how human-centric AI can be embedded in work design, governance structures, and skill development strategies. Longitudinal and cross-sectoral studies may additionally help examine how human-AI collaboration evolves over time in relation to e.g. skill development, cognitive load, trust dynamics, and potential overreliance on AI systems, as well as whether the identified human-centric aspects are transferable beyond manufacturing context.

5 Conclusion

The research question formulated for this study was: What human-centric goals, challenges, and enablers should be considered in the context of human-AI collaboration in manufacturing? Through a literature review, the findings identify trust, transparency, augmentation, human-AI capability complementarity, explainability, safety, and well-being as recurring human-centric goals. The goals contain inter-dependencies, such as trust, transparency, and explainability, forming a mutually reinforcing cluster, while ethics and well-being operate as meta-goals encompassing several others. Frequently reported challenges include job displacement, conflicting priorities, algorithmic bias,

and skill gaps and lack of training, while common enablers include human oversight, co-design, explainable interfaces, and task allocation. Overall, this study contributes a structured synthesis of human-centric aspects in human-AI collaboration in manufacturing and indicates the need for translating these aspects into actionable design and organizational practices.

Acknowledgments. The work in this paper has been conducted as part of the SkillAIbility project, and we gratefully acknowledge the supporting funding from the European Union's Horizon Europe research and innovation program under Grant Agreement No 101177783. The work is also part of Chalmers Production Area of Advance.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Salunkhe, O., Chávez, C.A.G., Wang, H., Syberfeldt, A., Romero, D., Stahre, J.: Developing code agents for robot programming: technical and managerial perspectives. In: Mizuyama, H., Morinaga, E., Nonaka, T., Kaihara, T., Von Cieminski, G., Romero, D. (eds.) Advances in Production Management Systems. Cyber-Physical-Human Production Systems: Human-AI Collaboration and beyond, pp. 134–147. Springer Nature, Switzerland, Cham (2026). https://doi.org/10.1007/978-3-032-03515-8_10
2. Kusters, R., et al.: Interdisciplinary research in artificial intelligence: challenges and opportunities. Front. Big Data. **3**, 577974 (2020). https://doi.org/10.3389/fdata.2020.577974
3. Hammad, M.Y., Rahamaddulla, S.R., Tamyez, P.F.M., Fauzi, M.A.: From Industry 4.0 to 5.0: Leveraging AI and IoT for Sustainable and Human-Centric Operations.
4. Xu, X., Lu, Y., Vogel-Heuser, B., Wang, L.: Industry 4.0 and industry 5.0—inception, conception and perception. J. Manuf. Syst. **61**, 530–535 (2021). https://doi.org/10.1016/j.jmsy.2021.10.006
5. Hermawati, S., Correa, R., Mohan, M., Lawson, G., Houghton, R.: Defining human-centricity in industry 5.0 and assessing the readiness of ergonomics/human factors communities in UK. Ergonomics. **68**, 857–876 (2025). https://doi.org/10.1080/00140139.2024.2343947
6. Müller, J.M.: Assessing the barriers to industry 4.0 implementation from a workers' perspective. IFAC-Pap. **52**, 2189–2194 (2019). https://doi.org/10.1016/j.ifacol.2019.11.530
7. Mentzas, G., Hribernik, K., Stahre, J., Romero, D., Soldatos, J.: Editorial: human-Centered artificial intelligence in industry 5.0. Front. Artif. Intell. **7**, 1429186 (2024). https://doi.org/10.3389/frai.2024.1429186
8. University of Maryland, Park, C., Shneiderman, B.: Human-Centered artificial intelligence: three fresh ideas. AIS Trans. Hum.-Comput. Interact. **109–124** (2020) https://doi.org/10.17705/1thci.00131
9. Leng, J., Mourtzis, D., Liu, Q., Xu, L., Kusiak, A., Pham, D.T.: Human-centric AI as a partner: models for augmenting creativity and expertise in industry 5.0. Int. J. Prod. Res. **1–9** (2026). https://doi.org/10.1080/00207543.2026.2662984
10. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and levels of human interaction with automation. IEEE Trans. Syst. Man Cybern. - Part Syst. Hum. **30**, 286–297 (2000). https://doi.org/10.1109/3468.844354
11. Endsley, M.R.: Automation and Situation Awareness.
12. O'Neill, T., McNeese, N., Barron, A., Schelble, B.: Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. Hum. Factors J. Hum. Factors Ergon. Soc. 64, 904–938 (2022). doi:https://doi.org/10.1177/0018720820960865.

13. Cimini, C., Pirola, F., Pinto, R., Cavalieri, S.: A human-in-the-loop manufacturing control architecture for the next generation of production systems. J. Manuf. Syst. **54**, 258–271 (2020). https://doi.org/10.1016/j.jmsy.2020.01.002
14. Mokhtarzadeh, M., Rodríguez-Echeverría, J., Semanjski, I., Gautama, S.: Hybrid intelligence failure analysis for industry 4.0: a literature review and future prospective. J. Intell. Manuf. **36**, 2309–2334 (2025). https://doi.org/10.1007/s10845-024-02376-5
15. Passalacqua, M., et al.: Human-centred AI in industry 5.0: a systematic review. Int. J. Prod. Res. **63**, 2638–2669 (2024). https://doi.org/10.1080/00207543.2024.2406021
16. Schmager, S., Pappas, I.O., Vassilakopoulou, P.: Understanding human-centred AI: a review of its defining elements and a research agenda. Behav. Inf. Technol. **44**, 3771–3810 (2025). https://doi.org/10.1080/0144929X.2024.2448719
17. Yan, J., et al.: Human-centric artificial intelligence towards industry 5.0: retrospect and prospect. J. Ind. Inf. Integr. **47**, 100903 (2025). https://doi.org/10.1016/j.jii.2025.100903
18. European Commission: Directorate general for research and innovation.: industry 5.0: towards a sustainable, human centric and resilient European industry. Publications Office, LU. (2021). https://doi.org/10.2777/308407
19. Muntean, M.H., Bocanet, V.I., Fratila, D.: AI enhanced quality assurance reporting of surface deviation data in optical 3D inspection. In: Mizuyama, H., Morinaga, E., Nonaka, T., Kaihara, T., Von Cieminski, G., Romero, D. (eds.) Advances in Production Management Systems. Cyber-Physical-Human Production Systems: Human-AI Collaboration and beyond, pp. 188–201. Springer Nature, Switzerland, Cham (2026). https://doi.org/10.1007/978-3-032-03538-7_14
20. Bitterling, C., Loske, D., Klumpp, M.: Improving visual inspection accuracy: explainable AI for human-centric quality control. IFAC-Pap. **59**, 1450–1455 (2025). https://doi.org/10.1016/j.ifacol.2025.09.244
21. Catti, P., Nikolakis, N., Alexopoulos, K.: A human-centric drift controller framework for adaptive and explainable quality control in manufacturing. In: In: 2025 IEEE 30th International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1–8. IEEE, Porto, Portugal (2025). https://doi.org/10.1109/ETFA65518.2025.11205586
22. Senoner, J., Schallmoser, S., Kratzwald, B., Feuerriegel, S., Netland, T.: Explainable AI improves task performance in human–AI collaboration. Sci. Rep. **14**, 31150 (2024). https://doi.org/10.1038/s41598-024-82501-9
23. Passalacqua, M., et al.: Practice with less AI makes perfect: partially automated AI during training leads to better worker motivation, engagement, and skill acquisition. Int. J. Human–Computer Interact. **41**, 2268–2288 (2025). https://doi.org/10.1080/10447318.2024.2319914
24. Galetić, V., Nottle, A.: Inherently Interpretable Knowledge Representation for a Trustworthy Artificially Intelligent Agent Teaming with Humans in Industrial Environments.
25. Hémono, P., Nait Chabane, A., Sahnoun, M.: Leveraging Digital Twin and Dynamic Scheduling for Enhanced Human-Robot Collaboration., https: //www.ssrn.com/abstract=5065955 (2024). https://doi.org/10.2139/ssrn.5065955
26. Chen, S.-C., Chen, H.-M., Chen, H.-K., Li, C.-L.: Multi-objective optimization in industry 5.0: human-centric AI integration for sustainable and intelligent manufacturing. PRO. **12**, 2723 (2024). https://doi.org/10.3390/pr12122723
27. Ramzan, F., Reforgiato Recupero, D.: A literature review on enhancing predictive maintenance in smart manufacturing industries: fostering human-technology collaboration and overcoming data scarcity limitations with advanced AI models. Oper. Res. Forum. **6**, 181 (2025). https://doi.org/10.1007/s43069-025-00584-0
28. Karapanagiotis, P., Kottagaha, W.M.,.K., Rovere, D., Bokhorst, J.A.C., Valdata, A., Emmanouilidis, C.: Enabling interoperable human–AI teaming for automation in construction and manufacturing via digital twins and sliding work sharing ontologies. J. Ind. Inf. Integr. **48**, 100962 (2025). https://doi.org/10.1016/j.jii.2025.100962

29. Sauer, C.R., Burggräf, P.: Hybrid intelligence – systematic approach and framework to determine the level of human-AI collaboration for production management use cases. Prod. Eng. **19**, 525–541 (2025). https://doi.org/10.1007/s11740-024-01326-7
30. Pan, Y., Zhang, W., Liao, B., Kong, M., Zhang, H., Fathollahi-Fard, A.M.: Smart outbound scheduling for textile warehousing: an integrated ALNS optimisation algorithm and LLM-based fuzzy decision-making framework. J. Eng. Des. **1–38** (2026). https://doi.org/10.1080/09544828.2025.2597722
31. Gholamizadeh, K., Zarei, E., Gualtieri, L., Marchi, M.D.: Advancing occupational and system safety in industry 5.0: effective HAZID, risk analysis frameworks, and human-AI interaction management. Saf. Sci. **184**, 106770 (2025). https://doi.org/10.1016/j.ssci.2024.106770
32. Merchán-Cruz, E.A., et al.: Trust by design: an ethical framework for collaborative intelligence Systems in Industry 5.0. Electronics. **14**, 1952 (2025). https://doi.org/10.3390/electronics14101952
33. Emmanouilidis, C., Zotelli, J., Hengel, K., Waschull, S., Bokhorst, J.: Co-creation design patterns for human-AI teaming in manufacturing and multi-domain decision-making. IFAC-Pap. **59**, 227–232 (2025). https://doi.org/10.1016/j.ifacol.2025.11.869
34. Żywiołek, J.: Trust-building in AI-human partnerships within industry 5.0. Syst. Saf. Hum. - Tech. Facil. - Environ. **6**, 89–98 (2024). https://doi.org/10.2478/czoto-2024-0011
35. Kadam, A.A., Peddinti, D.R., Kadam, S.A.: AI-driven mechanical manufacturing: bridging industry 4.0 and the human-centric vision of industry 5.0. In: Sinha, G.R., Fan, C.-P., Bajaj, V., Nisar, H., Ullo, S.L. (eds.) Integrating AI in Science, Management, and Technology, pp. 170–190. Springer Nature, Switzerland, Cham (2025). https://doi.org/10.1007/978-3-032-08260-2_13
36. Chigbu, B.I., Makapela, S.L.: AI in education, sustainability, and the future of work: an integrative review of industry 5.0, education 5.0, and work 5.0. J. Open Innov. Technol. Mark. Complex. **11, 100645** (2025). https://doi.org/10.1016/j.joitmc.2025.100645
37. Hoch, T., et al.: Teaming.AI: Enabling Human-AI Teaming Intelligence in Manufacturing.
38. Hartikainen, M., Spurava, G., Väänänen, K.: Human-AI collaboration in smart manufacturing: key concepts and framework for design. In: Lorig, F., et al. (eds.) Frontiers in Artificial Intelligence and Applications. IOS Press (2024). https://doi.org/10.3233/FAIA240192
39. Mamad, M., Chichi, O.: Towards a human-centred artificial intelligence in the age of industry 5.0: a cross-country analysis. In: In: 2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM), pp. 1–6. IEEE, Leeds, United Kingdom (2024). https://doi.org/10.1109/WINCOM62286.2024.10655038
40. Crnobrnja, J., Horvat, D., Romero, D., Ruppert, T., Marjanovic, U.: Analyzing the adoption of AI and human-centricity in Serbian manufacturing firms. In: Mizuyama, H., Morinaga, E., Nonaka, T., Kaihara, T., Von Cieminski, G., Romero, D. (eds.) Advances in Production Management Systems. Cyber-Physical-Human Production Systems: Human-AI Collaboration and beyond, pp. 78–88. Springer Nature, Switzerland, Cham (2026). https://doi.org/10.1007/978-3-032-03538-7_6
41. Tóth, A., Nagy, L., Kennedy, R., Bohuš, B., Abonyi, J., Ruppert, T.: The human-centric industry 5.0 collaboration architecture. MethodsX. **11**, 102260 (2023). https://doi.org/10.1016/j.mex.2023.102260
42. Brückner, A., Hein, P., Hein-Pensel, F., Mayan, J., Wölke, M.: Human-Centered HCI practices leading the path to industry 5.0: a systematic literature review. In: Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (eds.) HCI International 2023 Posters, pp. 3–15. Springer Nature, Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-35989-7_1

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

