Cognitive Computing and Internet of Things, Vol. 165, 2025, 187–197
https://doi.org/10.54941/ahfe1006102

Fairness in Designing Decision-Making Processes With Multi-Agent Systems and Human Factors

Gustavo Vieira[1,4], Herwig Zeiner[2], Lucas Paletta[2], Rui Fernandes[1], Higor Rosse[1], Khadija Sabiri[1,4], Juan F. de Paz[3], and José Barbosa[4,5]

[1]Collaborative Laboratory Mountains of Research, Intelligent Technologies, Bragança, Portugal
[2]Joanneum Research Forschungsgesellschaft mbH, Graz, Austria
[3]Universidad de Salamanca, Department of Computer Science and Automation, Salamanca, Spain
[4]Research Centre in Digitalization and Intelligent Robotics (CeDRI), Laboratório Associado para a Sustentabilidade e Tecnologia em Regiões de Montanha (SusTEC), Bragança, Portugal
[5]Instituto Politécnico de Bragança, Campus de Santa Apolónia, Bragança, Portugal

ABSTRACT

This paper explores the integration of Human Factors (HF) into Multi-Agent Systems (MAS) to enhance fairness in decision-making processes in Industry 5.0 environments. We contribute with a human-centred perspective in the development of MAS by integrating the physiological aspects of workers in the manufacturing industry. This culminates in the measurement of human resilience. This paper presents an automotive manufacturing environment where wearable sensors and AI-driven analytics assess workers' physiological and psychological stress levels to calculate a human resilience score. This score, along with worker preferences, supports a dynamic worker allocation algorithm based on MAS that adapts to production demands. Our approach embodies the Industry 5.0 vision of technologies that support adaptive, transparent and, above all, fair human management. The system uses advanced technologies to meet business goals and employee needs and to promote a more inclusive, supportive and people-centred work environment.

Keywords: Human factors, Multi-agent systems, Fairness

INTRODUCTION

Fairness in a societal context extends beyond algorithmic accuracy; it questions whether statistical fairness aligns with societal values and norms. Even if an algorithm achieves fairness in a technical sense, one must consider the impact in the social aspect (Starke et al., 2022). This broader perspective demands a holistic view, recognizing that fairness involves outcomes and decision-making processes. The debate around fairness concerns how these systems can integrate concepts of fairness that account for human relational dynamics. Fairness is a dynamic value, shaped by balancing different perspectives and ensuring that all relevant actors play a role in the decision process. Organizations should approach algorithmic fairness as an ongoing process rather than a fixed outcome.

Essential for ensuring impartial and just decision-making, fairness is a fundamental concept in democratic societies. In the context of Decision Support Systems (DSS), fairness becomes more complex, involving the need to address algorithmic biases, respect individual rights while maintaining transparency and accountability (Woitsch et al., 2024). Multi-Agent Systems (MAS) can address fairness in the real world. It represents stakeholders interacting in the real world, enabling a path to balance workers' resilience and preferences in a production setting.

Fairness must be conceived as a process that considers stakeholders' needs by leveraging Human Factors data rather than solely as an outcome that doesn't acknowledge how it was reached (Helberger et al., 2020). This involves actively engaging all relevant actors in the decision-making process. Human Factors (HF) data incorporated in MAS ensure that the system accounts for human capabilities, limitations, and behaviors, making the process more transparent, accountable, and aligned with societal values. This holistic approach leads to fairer outcomes and builds trust among stakeholders, as they can see how their needs and concerns are considered throughout the process.

This paper aims to provide an innovative approach to address fairness in decision-making processes, aligning MAS with an HF methodology that enables the measuring and balancing relevant parameters such as the resilience score to meet a fair decision outcome in a real-world scenario. In Section 2, related works are pointed out and briefly discussed. It describes how MAS enables fairness and promotes an architectural approach that facilitates addressing fairness in environments with many stakeholders. Next, in Section 3, Human Factors are introduced and integrated into the MAS framework. This section delves into how principles from Human Factors are applied to design systems that consider human capabilities and limitations for promoting fairness. Section 4 presents a car manufacturing case study that applies a proposed MAS solution for integrating fairness through Human Factors data in the presented use case. Finally, Section 5 concludes this work, highlighting future work.

RELATED WORK

Different Perspective of Fairness

Fair decision-making is highly relevant from different community perspectives. First, it is important to analyze the outcome of a decision-making process, especially for services that involve the distribution and allocation of resources. There are various techniques for implementing fairness, including methods based on distributed learning, distributed mean consensus, and game theory, as described in Starke et al. (2022), Landers and Behrend (2023) and Yin et al. (2023).

Second, fairness in AI-informed decision-making explores the relationship between people's perceptions of fairness and how decisions made by

AI systems are explained to them. The study suggests that providing explanations can increase trust in the fairness of AI-based decisions (Angerschmid et al., 2022). A study on perceptions of fairness and trust in decision-making examines the relationship between people's trust in automated decision systems and their understanding of how they work. The research shows that a lack of transparency can lead people to question the fairness of such systems (Schoeffler et al., 2022).

This study highlights that human input is fundamental to achieving a fair process (Uhde et al., 2020). Perceived fairness stems from the relevance and justification of reasons given for work shifts negotiated between peers. When employees actively participate in scheduling and understand the rationale behind any changes, it fosters a sense of transparency and trust within the team. This directly impacts workers' well-being and satisfaction in the workplace, as they feel their needs and contributions are valued. Higher levels of satisfaction can lead to increased motivation and better job performance. Moreover, involving employees in decision-making empowers them, promoting a collaborative environment. This underscores the importance of incorporating human input into decision-making processes that affect well-being and satisfaction, which in turn are key factors that influence the quality of work performed.

Finally, it is important to ensure proportional fairness and freedom from envy. These techniques are essential for achieving fairness in resource allocation (Jiang and Lu, 2019). Furthermore, in system dynamics and agent-based modeling simulations, fairness can be conceptualized by considering procedural fairness, which relates to the procedures that lead to outcomes, and distributive fairness, which relates to the perception of outcomes as fair or unfair. In general, we examine the field of algorithmic fairness and its goals. Using algorithms for automated decision-making can have unintended effects that lead to discrimination against certain groups (e.g., in the workload example). In this context, it is crucial to develop MAS that are not only accurate but also fair, considering the process that leads to it.

MAS ENABLING FAIRNESS

Fairness in MAS involves more than just designing algorithms; it requires understanding human fairness motivations and how these can be modeled and translated into a computational framework (de Jong et al., 2008). The challenge lies in capturing the complex nuances of human fairness, which often encompasses ethical, social, and emotional dimensions, and embedding these into systems where multiple agents interact. This involves ensuring that individual agents operate fairly and that their interactions lead to outcomes perceived as fair by humans.

The dynamics within MAS often mirror social dilemmas where collective interests clash with individual agents' goals. In such scenarios, the concept of fairness extends to understanding and balancing these conflicts. This requires an understanding of how agents can either contribute to or detract from overall fairness in emergent team behaviors. The development of cooperative multi-agent fairness thus reframes key questions to focus on whether agents, given incentives to collaborate, can learn to coordinate their actions effectively and fairly. However, pursuing fairness in MAS does not come without cost, especially as task difficulty increases. Empirical studies in cooperative multi-agent tasks suggest that while fairness may be relatively "inexpensive" in simpler scenarios — where agent skills are sufficiently high — it can become increasingly costly in more complex situations (Grupen et al., 2022).

Furthermore, the broader implications of fairness in decision support systems require a dual perspective that encompasses both algorithmic and societal views. On one hand, there is a need to develop algorithms capable of balancing different relevant decision factors within a defined context. On the other hand, it is crucial to consider whether the type of fairness achieved by these algorithms aligns with societal values (Angerschmid et al., 2022; de Jong et al., 2008). This distinction highlights the importance of not only designing decision support systems that are fair in a statistical sense but also ensuring that these systems contribute to a form of fairness that is meaningful and desirable within the societal context. This dual perspective underscores the ongoing dialogue and necessary adjustments in how fairness is conceptualized and implemented in MAS and broader automated systems.

MAS reproduce these behaviors through the development of descriptive models of human fairness that can be further explored to enhance the decision-making capabilities of these systems to reflect fairness akin to human values. In the next steps, we aim to explore fairness aspects in MAS to provide an approach that aligns with human values, which, in fact, are desired in decision support systems.

Human Factors Promoting Fairness in the Context of Industry 5.0

Including human factors in the design and implementation of fairness promotes a human-centered perspective, encompassing a range of considerations from psychological and cognitive aspects. This approach highlights the importance of understanding and incorporating the human perspective into systems, with a specific focus on human resilience as a fundamental dimension for well-being.

When discussing reactions to stress in the workplace, it is important to mention the individual characteristics of resilience. People with a high level of resilience are considered to be better able to cope with stressful work demands than other employees. Resilience positively correlates with job satisfaction and negatively with the likelihood of absenteeism, burnout, and, eventually, productivity losses (Paletta et al., 2024). Resilience can be conceptualized as a trait but is also generally viewed as a process or capability that can be developed in a temporal context. As part of this understanding of the process, some companies have begun to implement measures to increase the resilience of their employees. A systematic review (Blandino, 2023) found that the duration of these interventions can vary from a single session to a 12-week program and that a variety of techniques, such as skills-based coaching, mindfulness, and compassion-based practices, cognitive behavioral techniques, such as energy management and relaxation training, are used.

In this work, a combination of stationary and wearable sensors and AI-driven analytics allows for assessing a worker's physiological and psychological stress levels, offering insights into their resilience. A network of local wearable sensors gathers bio-signals from devices worn by the workers. This system then produces higher-level abstractions of human factors, providing a comprehensive view of the worker's condition.

The resilience score reflects a worker's capacity to handle physiological or cognitive-emotional strain without experiencing negative long-term effects. Workers with high resilience scores can perform in more stressful environments and handle more demanding tasks compared to those with lower scores. This score integrates a measure of the physiological strain of a worker, estimated through the Physiological Strain-Index (PSI), which is calculated based on measured skin temperature, heat flux (via a biosensor on the chest), and heart rate (monitored by a biosensor shirt or intelligent armband) as displayed in Figure 1.

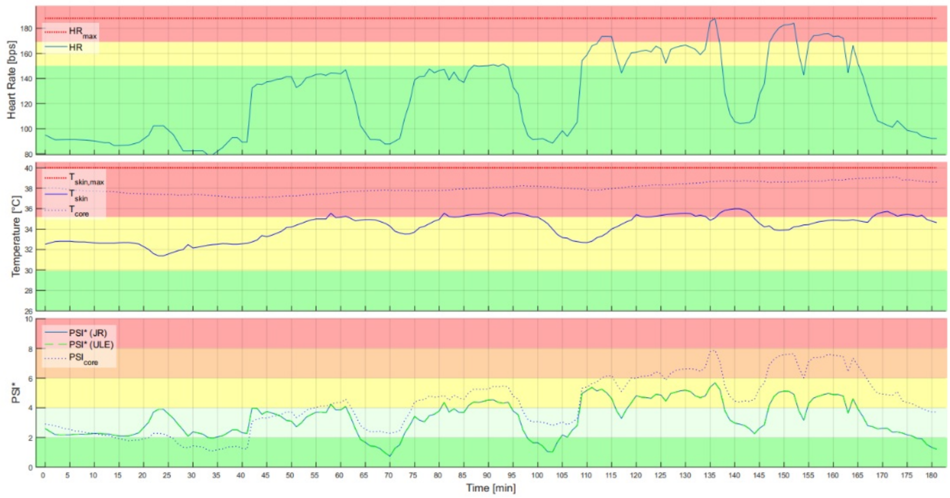

Figure 1: Human factors data: heart rate, skin and core body temperature, and physiological strain-index.

For the resilience calculation, a function that aggregates PSI values over time is used, offering an indication of the intensive fatigue experienced by the worker that could affect their mid or long-term resilience. The PSI ranges from 0 (no strain) to 10 (maximum individual strain) and is divided into five zones: [0–2], (2–4], (4–6], (6–8], and (8–10]. The weighted sum of time spent in each PSI range is normalized to an eight-hour shift, and the resilience score for the day, $\overline{R}(\mathrm{d})$ in Equation 1, is inversely proportional, ranging from 0 to 1. A heuristic estimate of future capacity is derived from the resilience scores of the previous 20 working days. For each new working day, a fresh resilience score is calculated using a linear combination of the last 20 scores. The overall resilience score $\overline{R}(\mathrm{d})$ is ultimately determined by a weighted average of the scores from the most recent 20 days (Vieira et al., 2023).

$$\overline{R}(\mathrm{d}) = \left(\sum\nolimits_{n=1}^{20} w_{-n}\, R_{d-n} \right) \tag{1}$$

CAR PART PRODUCTION SYSTEM

In the manufacturing industry, workload balance is relevant to ensure that production goals are met while keeping workers' well-being in mind. This problem occurs in every factory where the production team must plan the upcoming week's workload. The process initiates by an event, such as a new production schedule or a shift in production requirements due to increased orders or changes in the product line.

Use Case Description

In this study, a factory that produces various automotive components is considered. Three production lines composed of six and four working stations receive workers to execute the production of car parts in a sequence. Each order has a specific number of steps that require workers with specific capabilities. One worker is allocated to each workstation based on their skill set and the requirements of each production step. The production lines operate sequentially, meaning that a part progresses from one workstation to the next as each step is completed. Each week, the production manager receives a schedule outlining the quantity of parts needed. This event triggers the need to allocate or reallocate workers to different tasks on the production floor to meet the operative requirements, namely the throughput and workload balance.

Following the strategy to fulfill a fair distribution of load among human operators, the workload balance process begins with assessing two key parameters: worker availability and medical condition. Worker availability regards the presence of the worker in the required working shift. Workers might be absent e.g. due to unforeseen events or sickness. The medical condition refers to the capability of a worker to perform the task, considering it is a stevedoring process. Based on this information, the system can dismiss workers from the possible choices to allocate a worker in the upcoming tasks of the production order list. Next, the system checks each worker's capabilities by accessing a database where this information is located. With this data, the production manager can assign tasks that best match their qualifications.

The allocation of workers proceeds in two stages. The first stage involves assigning workers to production lines based on their general skills and the production requirements. In the second stage, the system recommends worker allocation to balance workload and worker capability with production demands. For example, if a line requires a particular skill that only a few operators have, those operators are strategically placed in roles that demand their expertise. However, the worker's well-being must not be neglected.

The decision support system is essential in both stages, as it analyzes production deadlines, worker availability, skill requirements, and worker's human factors to allocate them effectively. This system optimizes production efficiency and considers the human as a center point of the production process, creating a balanced workload distribution that supports both production goals and worker well-being.

Multi-Agent Fairness in a Real-World Use Case

The use case consists of a workload balance scenario with a systematic process for optimizing worker allocation in a car part manufacturing setting (see Figure 2). The scenario is initiated by weekly production planning. This process includes parallel tasks that assess production requirements, determining the number of workers needed to meet orders. Before assigning workers, their availability and physical condition to work with the particular load lifting of car parts, named as medical condition, are verified through a database. The decision-making process, managed by the plant's manager of the production section, involves periodic checks to ensure workers meet availability and medical criteria for specific production lines and geometries according to the NIOSH (National Institute for Occupational Safety and Health) index.

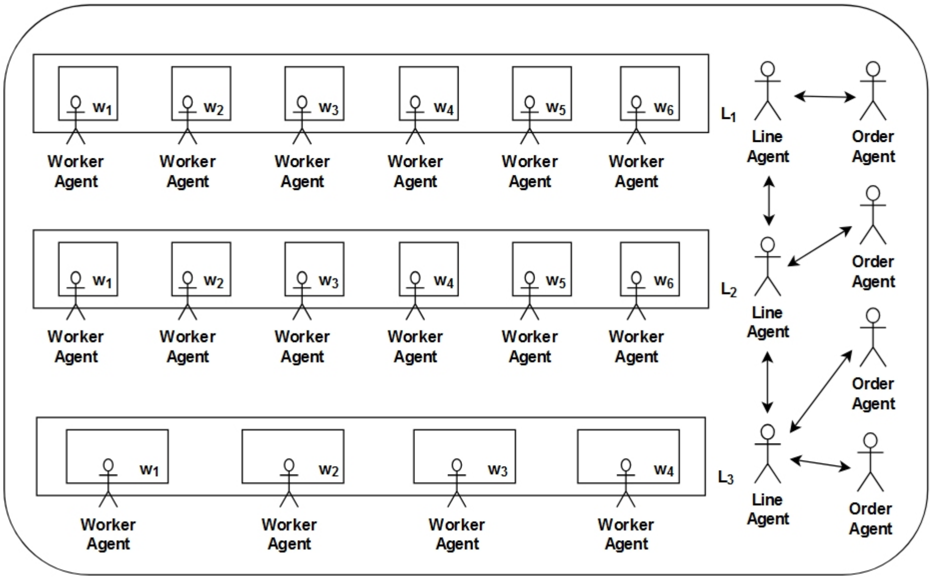

Figure 2: Multi-agents in a car manufacturing use case.

The Industry 5.0 concerns how human integration takes part in the decision-making process. This use case scenario presents the opportunity to improve its process through a human-centric approach that leverages human participation, representation, and relevance in a decision process that originally missed broad accountability to human needs. Fairness in decision-making can be achieved through the exploration of how the production requirements can be dealt with human nature needs. In addition to tracking human physiological data, an essential aspect of this case is the digital representation of interactions among individual requirements. The possibility of coordination and negotiation (see Figure 3) of stakeholders in a digital realm allows for an improved decision since these parameters of interest can be measured, tracked, and assigned in a goal-oriented way. Hence, these parameters of interest regarding conflicting goals can be managed to promote a decision that accounts for human needs and production requirements. It is important to observe that leveraging these aspects through digital approaches fosters a more inclusive and supportive work environment, embodying the principles of Industry 5.0 that bring added value to the decision process.

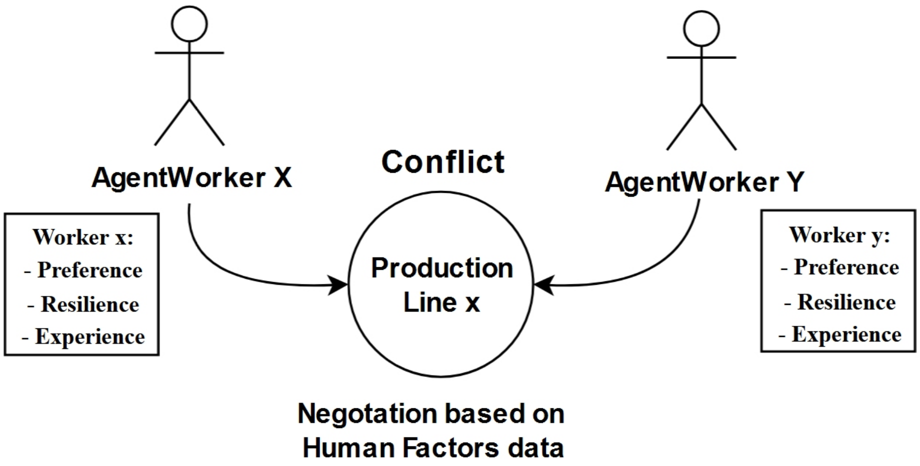

Figure 3: Agent interaction.

In this worker allocation scenario, the decision-making process involves criteria that ensure operational efficiency and worker well-being. Initially, the production requirements drive the allocation, determining the number of workers needed based on the demands of the car parts manufacturing process. Key factors include production priorities, worker availability, and compliance with health and safety standards, such as those outlined by the NIOSH index. Worker assignments are evaluated based on their ability to meet these requirements while simultaneously considering their conditions, such as medical limitations. The decisions are also influenced by real-time checks on worker availability and their suitability for specific tasks, ensuring that human factors like fatigue or medical constraints are respected.

Fairness in the decision-making process is assessed by considering operational and human-centric metrics. These include the equitable workload distribution among workers, including worker preferences and individual physiological needs, and compliance with health and safety regulations (NIOSH index). The process also involves tracking individual worker resilience and ensuring that high-demand tasks are rotated among employees to avoid physical or mental strain. In addition, fairness is promoted by digitally measuring and incorporating worker preferences and capabilities into the worker allocation process, ensuring that no worker is disproportionately assigned to difficult or repetitive tasks without consideration of their physiological capabilities or preferences.

The algorithm behind the decision-making process operates in stages to dynamically allocate workers to tasks of which the use case UML can be seen in Figure 4. Initially, it receives inputs from two primary data sources: the production line requirements and worker information (availability and medical condition). Then, it will check the worker's resilience, which reflects how fit the worker is to perform the task in each production line, ranging from a value of 0 to 1. Additionally, the production priority and due date for the products are also checked. If the worker registered in the pool of workers is not available or does not sustain the medical condition to perform the task in the given production line, it is immediately dismissed from the allocation process.

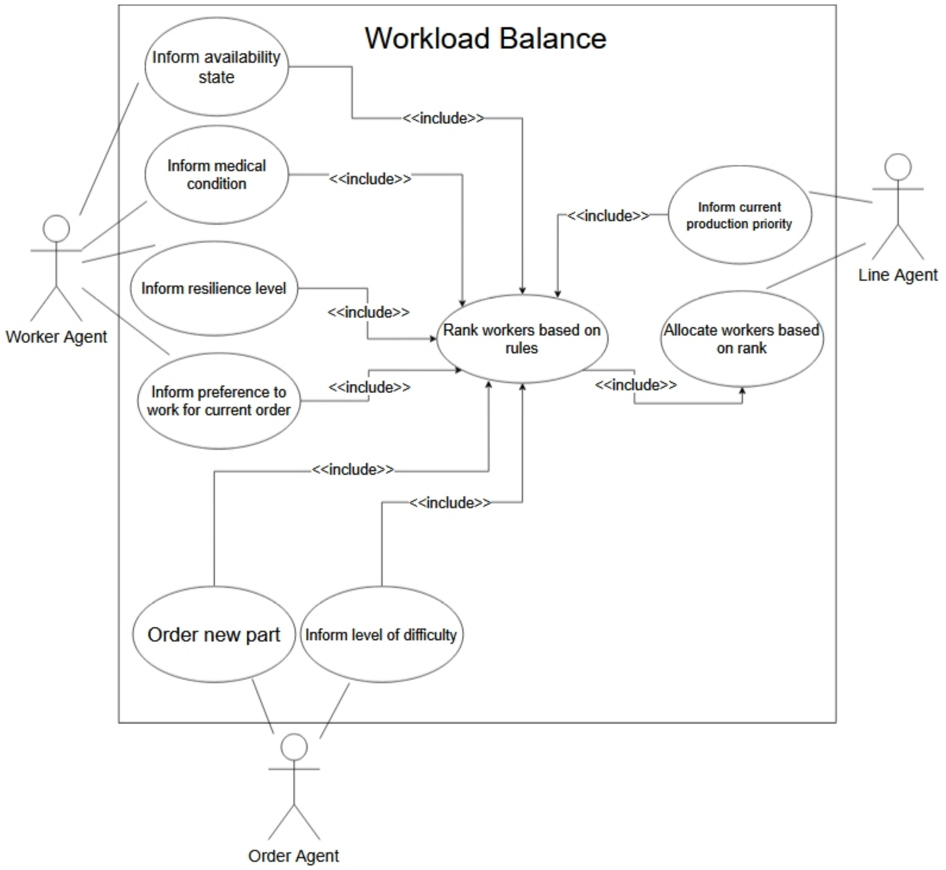

Figure 4: Multi-agent use case UML.

Then, a score is calculated for each worker based on two parameters: worker resilience and worker preference. For this prototype, we consider three lines, among which the workers select a preference level for each, ranging from 0 to 1. Based on that, a score is calculated, and its value is used in the agent interaction to negotiate the allocation of suitable workers to the production lines. If the system dynamically adapts to changing production conditions whenever requiring different levels of worker resilience for different production requirements and workers' individual preferences, a balanced and efficient work environment that supports both operational goals and the satisfaction of workers will be promoted.

CONCLUSION

This paper provides an insight into the use of human aspects in decision-making. These include human physiological factors, particularly the concept of resilience, and human input on preference in task execution. Furthermore, the concept of fairness in the context of Multi-Agent Systems has been analyzed to understand how it can be used to make decisions and implemented, attaining a fair decision-making process applied to a relevant industrial scenario. When these factors are analyzed from both human and technical perspectives, the paper provides valuable insights that can help developers and practitioners improve their decision-support systems. This comprehensive understanding of the challenges and requirements of real-world scenarios facilitates the development of tailored solutions that meet the demanding requirements of manufacturing.

Calculating a resilience score based on physiological data and incorporating worker preferences enables the system to adapt to changing production conditions while promoting worker satisfaction and well-being. This approach embodies Industry 5.0's vision of technology supporting adaptive, transparent, and fair human aspects. Integrating MAS and HF principles demonstrates how advanced technologies can be harnessed to meet organizational goals and worker needs.

Future work will involve applying the experimented concept on a larger scale while establishing additional feedback mechanisms where workers and managers can provide input on the allocation process outcomes to ensure continuous improvement and alignment with human-centric principles.

ACKNOWLEDGMENT

This work has been supported by the FAIRWork project (www.fairwork-project.eu) and has been funded by the European Commission's Horizon Europe Programme under contract number 101069499. This paper expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained in this paper. The authors express their gratitude to the project consortium for their input in writing this paper, specifically the industrial partners Stellantis/CRF and Flex, for their valuable time during the requirements elicitation phase.

REFERENCES

Angerschmid, A., Zhou, J., Theuermann, K., Chen, F., & Holzinger, A. (2022). Fairness and Explanation in AI-Informed Decision Making. Machine Learning and Knowledge Extraction, 4(2), 556–579. https://doi.org/10.3390/make4020026

Blandino, G. (2023). How to Measure Stress in Smart and Intelligent Manufacturing Systems: A Systematic Review. Systems, 11(4), 167. https://doi.org/10.3390/systems11040167

de Jong, S., Tuyls, K., & Verbeeck, K. (2008). Artificial agents learning human fairness. Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2, 863–870. Presented at the Estoril, Portugal. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.

Grupen, N. A., Selman, B., & Lee, D. D. (2022). Cooperative Multi-Agent Fairness and Equivariant Policies. arXiv [Cs. AI], http://arxiv.org/abs/2106.05727.

Helberger, N., Araujo, T., & de Vreese, C. H. (2020). Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making. *Computer Law & Security Review*, Volume 39, p. 105456.

Jiang, J., & Lu, Z. (2019). Learning Fairness in Multi-Agent Systems. arXiv [Cs. LG], http://arxiv.org/abs/1910.14472.

Landers, R. N., & Behrend, T. S. (2023). Auditing the AI auditors: A framework for evaluating fairness and bias in high stakes AI predictive models. American Psychologist, 78(1), 36–49. https://doi.org/10.1037/amp0000972

Paletta, L., Schneeberger, M., Haid, F., Mosbacher, J., Pszeida, M., Tschuden, J., & Zeiner, H. (2024). Resilience scores from wearable biosignal sensors for decision support of worker allocation in production. In Cognitive Computing and Internet of Things (Vol. 124, pp. 155–165), AHFE open access.

Schoeffer, J., Kuehl, N., & Machowski, Y. (2022). "There Is Not Enough Information": On the Effects of Explanations on Perceptions of Informational Fairness and Trustworthiness in Automated Decision-Making. Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, 1616–1628, doi: 10.1145/3531146.3533218.

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. Big Data & Society, 9(2). https://doi.org/10.1177/20539517221115189

Uhde, A., Schlicker, N., Wallach, D. P., & Hassenzahl, M. (2020). Fairness and Decision-making in Collaborative Shift Scheduling Systems. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–13. Presented at the Honolulu, HI, USA. doi:10.1145/3313831.3376656.

Vieira, G., Zeiner, H., Paletta, L., Scheneeberger, M., Rosse, H., Fernandes, R., Olbrych, S., Nasuta, A., Dienstl, M., Muck, C., & Chevuri, R. (2023). Initial dai-dss prototype. https://fairwork-project.eu/deliverables/D4.2_Initial%20DAI-DSS%20Prototype%20v1.0a-preliminary.pdf

Woitsch, R., Muck, C., Utz, W., & Zeiner, H. (2024). Enable flexibilisation in fairwork's democratic ai-based decision support system by applying conceptual models using adoxx. Complex Systems Informatics and Modeling Quarterly, *(38)*, 27–53.

Yin, T., Raab, R., Liu, M., & Liu, Y. (2023). Long-Term Fairness with Unknown Dynamics, http://arxiv.org/abs/2304.09362.