sciendo

Human centricity and technology advances research in social sciences: A literature insight on artificial intelligence

Radu-Ioan Popa [1*]

[1] *"Lucian Blaga" University of Sibiu, Faculty of Social Sciences and Humanities, Lucian Blaga Street, no. 2A, 550169, Sibiu, Romania*

KEYWORDS

Artificial intelligence
Human centricity
Social perspective

ABSTRACT

Evolving technology and human-machine interaction are two major stances that question human centricity as a core dimension for the future placement of human variables and social perspectives inside intelligent systems. The present paper explores present challenges and an up to date research synthesis concerning social sciences dimensions, theories, concepts and implementation with regard to the study of human centred artificial intelligence. The main scope consists of drawing several red lines for future research and theoretical social insights on artificial intelligence at work. Following a scholarly literature insight and several reference works, the study explores major topics of interest analysed so far and proposes new directions for the evaluation of human variables and machine learning at work in an organizational setting. Results showed there is a vital need for expanding research from human centricity point of view on the present matter, where studies are still scarce and heterogenous. Moreover, concept clarification and theoretical explanations are yet to be sufficient in explaining people interactions and management of artificial intelligence at the workplace. In conclusion human empowerment at work through ethical and person oriented intelligent systems in a given organizational context can ensure a fair and efficient development of people's skills, professional objectives or level of performance while promoting human values, ethical principles and preserving the well-being. The paper presents also several practical implications of social sciences domain on artificial intelligence exploration and a few future directions for research.

* Contact address: radu.popa@ulbsibiu.ro (R. Popa)

Introduction

Constant new technological development in a "fast forward" society determined gradually the appearance of a gap between the sophisticated and speedy evolution of tech and user oriented strategy, where quite often we wonder whether the user nowadays has to adapt to it and not the either way around it.

One challenge that resides in the analysis of human centricity with regard to modern technologies refers to the scarcity of scholarly articles which address the issue in terms of social science perspective on one hand. On the other hand, there is a wide range of research articles that focus mainly on the engineering process, product development, new artificial intelligence (AI) based platforms and so on. The social sciences perspectives have yet to offer a full comprehension of the phenomena, with a few research articles that address the issue of social variables in such settings.

The present article underlines several up to date implications on human centricity when confronted with technological progress and the AI subject from a social science point of view. The main scope is to identify certain "red lines" concerning the research of AI and human variables, taking into consideration debating stances and existing literature. The research apparatus has long been exploring, developing and implementing AI systems, many a time leaving aside the human factor, human variables, values, human interaction and not just user service and parameters.

There is already a new concept alongside AI entitled HAI (Human Artificial Intelligence) or HCAI (Human Centered Artificial Intelligence), which designates the need for future technology to be developed as human centred, considering human views first, human values, conditions and context (Shneiderman, 2020a; 2020b; 2021; Yang et al. 2021). Apart from tools, patterns and platforms, there is a pressing need to understand AI in human terms, regulations and policies knowing that sectors like industry, research, education, medicine, media, and communication already work with such technology or develop in this direction. Improving human condition, the welfare, performance, and other aspects of society require a great attention

towards clear policies, respecting rights and countering inequality. In this view the rapid development of technology and AI base systems even if they by principle address the human needs and are built around a positive discourse of human condition improvement, there is a long way for implementing and testing its outcomes on a validity feature. If it were to make a simple analogy with calibrating a test validation, AI and human centricity formula follow the same rule, making sure that the technology does what it says and plans to implement.

In recent years, some of the scholarly articles encompassed a few socio-technical variables in the research of human-machine learning and usage. Most works focus on ethical challenges, distributive justice, discrimination, exclusion, transparency, communication and perception (Abdelaal 2021; Boada, Maestre, and Genís 2021; Giermindl et al. 2022; Dwivedi et al. 2023; Heyder, Passlack, and Posegga 2023). From the ethics perspective for example, several authors observed there are certain research gaps in exploring the topic when it comes to human-AI interaction taking into account complexity, limitations, restrictive character, duty ethics focus, information systems technical stance which rule out quite frequent the human actor and behavior (Te'eni et al. 2019; Amershi et al. 2019; Siau and Wang 2020; Berente et al. 2021; Islam and Greenwood 2021; Mirbabaie et al. 2022). Moreover, Heyder, Passlack and Posega (2023) described and explained in a conceptual framework the major notions from the scholar database that should accompany AI ethics research and development in vital keys such as: transparency, accountability, privacy, security, justice, fairness, benefits, sustainability, responsibility, autonomy and humanity. All these principles may shape the sociotechnical system theory and sociomateriality approach, giving more reason to duty ethics and virtue ethics. At the same time, the focus should address alike the social entity stance and sociomaterial common ground, beyond just material patterns. In another work, Restrepo Amariles and Baquero (2023) claimed the need for a co-evolutionary approach between human users and AI systems, while favouring meaningful human control and a human centric rule. The same authors draw attention that at the moment

most regulations focus just on marketing driven objectives, leaving aside fundamental rights central role, they neglect accountability in other phases of the AI life cycle (not just design phase), do not offer a complete legal explainability beyond technicality and do not test systems algorithms in terms of proportionality, necessity and adequacy (Restrepo Amariles and Baquero 2023).

Beyond the ethical ground, research on AI and the human factor has been abundant on the positive technical outcomes so far, from medical devices, industry manufacturing processes, transportation mapping, supply chain management, data analysis, aerospace industry, communication and media sectors. Yet undoubtfully the benefits, there is a growing concern when it comes to the human variable, which in some cases just fills the consumer role while in other cases missing at all from the formula. Studying the human presence inside the machine interaction and development must comprise far more detailed and centred human dimensions.

Theoretical social insights on AI research

One perspective that unfolds in recent years in human – machine interaction research, from a social science point of view, follows people attitude exploration towards various technology products, from virtual reality (VR), mixed reality (MR), augmented reality (AR) and finishing with AI and HAI. Attitudes of course reflect the user perception when either dealing with intelligent tech, working with it directly or indirectly, knowing about it and so on. On one hand, several studies identified negative perceptions and attitudes such as anxiety or resistance through low levels of engagement or trust (Li and Huang 2020; Braganza et al. 2021; Del Giudice et al. 2023) while others refer to a more neutral stance which leads towards expectations of sociality or humanity inside the interaction, curiosity or gain model (Hu, Lu and Gong 2021).

Perception of efficiency, alongside expected results and performance, task easiness and process simplicity may all generate attitudes at work that

lead to a certain reaction of acceptance, understanding or resistance and avoidance behaviour when working with AI systems. Moreover, inside the organization, employees feel the need to test first the tangible outcomes of any new technology that comes in hand at times when the risk of being jobless or replaced by it is closer to a market reality. In this regard, organizational culture may determine as well specific attitudes and behaviours towards smart tech, consolidating fears of the unknown or at the opposite clarifying them. In order to give an example, it can be stated that organizational myths underline the best such reactions. A "fear of being fired when technology takes in" myth existing among employees at a certain time in the organization may well fuel a high level of resistance towards the introduction of new assistive training stations that use Mixed Reality in their platforms, making the employee refuse, stall, avoid, block, miss, sabotage the training process, steps, performance and real objectives for evaluation.

Chowdhury et al. (2023) summarized several AI capabilities from the scholarly literature (e.g. data sources, technical infrastructure, transparency, time requirement, technical skills, business skills, leadership, organizational culture, change capacity, knowledge, governance and regulations etc.) that need more insight and research while studying also the direct impact on the human workforce (e.g. job design and features, autonomy, trust, psychological variables, motivation, job satisfaction, team development, creativity, innovation and career management) which solicits an even greater attention. The same authors place the outcomes in terms of process efficiency, data driven decision making, product or service innovation, customer satisfaction, employee productivity, sustainable business performance and brand image, which describe the numerous positive potential results in AI implementation and development (Chowdhury et al. 2023). Beyond these sectors already in place, the human centricity approach solicits more interest in studying the human factor at work in relation with AI systems and process, the individual or the person for which AI was created and developed. In another study, Yang et al. (2021) discussed the implications of AI, beyond individuals, at societal level where the effect ranges between enabling and

inhibiting society development. The same authors mention among those barriers the following (Yang et al. 2021): an increase in prime resources in order to ensure energy consumption for such systems; the presence of errors, biases and misuse of algorithms which may generate discrimination, inequality, pattern thinking, prejudice and stereotyping; hard access for AI development in the case of low and middle income countries; job inequality due to high AI knowledge requirements. Even if, at product and activity sector level, AI will drive the change, from smart home, education, medicine and industry to agriculture, energy, economy, smart cities, transportation, there is an incomplete understanding for personal outcomes and social implications at this stage concerning individuals, groups, communities and societies. Just a few barriers mentioned above by Yang et al. (2021) constitute a warning sign and a vital request to address these issues in future works of research and patterns of implementation for a human centred philosophy and view.

Hua et al. (2024) focused on the acceptance dimension when working with AI in the healthcare system, observing that employees are influenced by key factors such as system use (containing work value expectation, level of burden and self-efficacy), socio-organizational-cultural dimensions (containing workflow integration, social influence, organizational readiness, threat perception and ethics) and user features (containing tech openness, AI literacy, system understanding and level of trust). According to the authors, all these key factors shape the level of acceptance in working with AI systems in the health care sector (Hua et al. 2024).

For example, an employee that works with AI systems or has contact with it, will be likely to adopt an attitude based on the personal level of understanding for the technology, how open the person is towards this type of system, how much trust it is allocated towards its functions and outcomes and of course the level of knowledge and experience of the individual when working with AI. Moreover, the employee is present not just as a user profile in this formula but also as a centred dimension in work processes where the person will treat AI in terms of integration, organization adaptation,

perception of threat, following ethics and level of influence, expectation of performance, efficiency, task difficulty and values. Beyond the health care system, these guidelines may apply in any organizational setting when working with AI systems, given the fact that perceptions and attitudes are formed following basic social rules. In another study, Kandul et al. (2023) stated that in some specific experimental situations, participants underestimated the AI predictability in terms of performance and naturally overestimated the human prediction skills, opening an interesting debate topic over the human abilities to exercise control efficiently over AI in certain environments. Human centricity from this point of view should explore also human control design, efficiency objectives, level of trust and predictability for a better understanding and application of management norms inside AI control and regulations. Perception formation and impression in the human setting require the same social rules when confronting AI systems. When working with a new device that can give a reply and analyse human interaction, the human user will apply naturally a subjective evaluation, shaping the personal impression under the rule of expectancy, results, performance, efficiency, fairness and so on. At this point people have the natural tendency to perceive the AI in terms of equality norms and so to build their views, develop behaviours and manifest attitudes according to this social "equation". In an interesting experiment, Li, Chu and Xu (2023), observed that in a communication mediated model with AI, human users were influenced by the level of fairness displayed by the system alongside a certain degree of team awareness or entitativity, which may shape human impression, acceptance or avoidance. The same authors noted that there is a positive correlation between equality norms and organized perceptions in human-AI interaction, drawing attention that people have a tendency to evaluate the AI behaviour in human behavioural norms, rules and conditions (Li, Chu, and Xu 2023).

Under the terms of social interaction, machine learning and interaction are naturally bound to the same perception mechanism from the user point of view, which she or he apply in daily interactions with other fellow

individuals. Human centricity theory should explore more these principles for the future, knowing that the AI environment follows a different development path and the rules of social interaction which people apply in their daily lives between themselves. This might not function when managing AI systems. If so far the analysis focused on observed perception and attitude formation in various organizational settings when working or experimenting with AI technology from an individual stance, the group or team level should be equally of interest, given their impact on decision making, company performance and functioning.

In this regard the scholarly literature and research underlined the impact of AI environments on team life, performance and perception beyond individual level. Moreover, on a broader organizational scale, companies need to address more frequently nowadays the subject of virtual teams, teleworking and digital work space which are accompanied yet in early times by AI agents, platforms, smart devices, beyond simple computational software and analysis. The mixed teams, humans and AI actors have been in place now for some time inside the work area, studies trying to explore the perception level here as well, performance improvement, roles exchange and motivation. From the medical sector, engineering and constructions towards communication and entertainment, mixed human AI teams have already experienced resource sharing and work interactions. Since the theory behind social functioning of human teams is well known, the question that arises is how to deal with a human AI hybrid team? In one study, results showed that if AI team agents influence level decreases, it can lead to human performance over time, giving way for the fact that human team members will enjoy working with AI team members as long as the latter will provide chances for them to learn, improve and develop (Flathmann et al. 2023). In their study the same authors observed that human users inside the mixed team perceive and analyse in different ways the level of social influence and the level of trust in AI team members based on personal motivations, objective correspondence with the virtual assistant, personal preference for learning by watching or by

doing, taking action for winning or just for simple entertainment, supporting healthy beliefs, principles and views over teamwork (Flathmann et al. 2023).

Flathmann et al. (2023) also mentioned the importance of real life environment, context and risk that the respondents consider when adapting perceptions, motivations and variance of working with AI team members. Vishwarupe et al. (2022) also stated the importance of distinguishing and selecting momentum of when to trust or not the AI decision or to refer to human judgement instead, inside the human AI confluence especially when the intelligent systems can falter. The authors also created and presented a confluence framework with user centred design consisting of several dimensions such as (Vishwarupe et al. 2022): involving various stakeholders at every stage of the process, establishing a clear distinction between stipulated roles, regulating data and user privacy, coordinating virtual reality, mixed reality and augmented reality in an intuitive design, attributing explainable and interpretable character to intelligent systems while paying attention to context and environmental factors, beyond individual and AI levels. The authors suggest that beyond the classical view which used to aim only at understanding the use context, user demands, design patterns, evaluation and implementation, the new confluence model should also comprise the setting beyond the simple application, determine other stakeholders to participate, use various design patterns, more complex evaluation and solution implementation alongside constant feedback (Vishwarupe et al. 2022).

From this perspective the need to involve more social actors in not only creating and developing AI systems but also monitoring constantly, offering feedback and evaluation every step on the way is quite the challenge but also a vital demand. The team term and AI collaboration can be defined at work also from the multiple stances of human centred AI teams which offer a more appropriate and in-depth understanding, perception and interpretation for developing the next generation human-machine interaction. Beyond mixed teams that use human and AI agents for fulfilling various work tasks and the team members perceptions, the social perspective points out a more inclusive

and comprehensive framework when integrating AI systems in the human centred paradigm. Future research on social grounds concerning human centricity and AI related topics must consider extending perspectives beyond traditional team dynamics processes and functional human-AI designs algorithms. Taking into consideration all social actors that must participate and collaborate, intelligent systems developers, final users, regulation institutions and policy makers, AI standard agents and many more, it can be stated that team-AI related view rises at a macro level.

HCAI trends and variables require more conceptual clarity and terminology consistency, in order to achieve more solid approaches and common ground when it comes to research human AI interaction. In another study, Bingley et al. (2023) found several differences between developers and users perceptions on AI usage, where the first category of respondents focus more on ethical, privacy and security demands and less on the social impact of AI, while the latter value more human centricity and how they are understood by intelligent systems. More specific, AI developers prioritize themes such as functionality, ethics and data processing while users top social impact, functionality and understanding when working with intelligent systems (Bingley et al. 2023). The study results, as seen in other studies, outlined the importance of making AI work for the user, clarifying the impact on social actors, understanding user needs and making the person feeling understood, extending the policies beyond essential and towards HCAI objectives, conducting more research on people awareness (Bingley et al. 2023). Sociology, social psychology and other disciplines from the same domain can ensure future studies that will allow the understanding of human variables and needs put into practice with regard to AI, making concepts and theories operational, optimizing teams and individuals in their achievements and goals.

HCAI inside team management must bring a new user interface and approach, in which human led principles are core drives for the creation of an AI system that offers support for human development, training, social needs, towards a positive outcome for people.

A third and final research insight on AI from a social perspective, beyond individual perception or team level interaction consists of activity sectors outcome review and methodology testing. Some of the initial studies focused on human variables at work in various sectors, with preference towards the medical organization environment, engineer industry, transportation, communication and education areas.

For example, Zhu, Fan and Zhang (2019) presented that major research in intelligent systems still circulates mostly in fields such as automotive industry (41.30%), aerospace sector (19.57%), industrial plants (19.57%), energy sector (10.87%) and military industry (8.70%). Sharma, Yadav and Chopra (2020) underlined a more intense presence of AI inside the economic and financial sector, transportation and environmental sustainability, with information and communication domains, policy making and law, healthcare still underrepresented. On a more specific note, other authors showed that medical sector is catching up with AI being represented by machine learning, artificial neural network, natural language processing in clinical devices, support vector machines and heuristics analysis platforms build, developed and implemented for diagnosis, personalized treatment and drug development, surgery, radiology, hospital administration and individual health records, cardiology and so on (Haleem, Javaid and Khan 2019).

At the moment the scholarly literature solicits more studies with focus on the social variables inside the activity domains mentioned before, and not just from singular, isolated social concept point of view but towards a more integrated approach, extending beyond simple classical human factors and plunging into deep understanding of human personality, motivations, behaviours, cognitive processes and many more when interacting with AI at the workplace. The process of integrating the human variable in a human centred paradigm and philosophy in AI management requires not only research and conceptualization but a certain dynamic ability for adaptation in parallel with the fast track of AI technology progress. In other words where tech evolution increases dramatically, social theory, social practice and regulation must follow close by and if possible be a step ahead to ensure that

human needs, interests and well-being are not shadowed by a technology only and overly driven development. At present, the balance between human centricity and techno-centricity resides around the latter, but with the evolution of socio-technical systems and a greater attention to social and human variables, the new industry must shape and adapt accordingly. The transition from simple AI use to HCAI implementation is not simple due to the complexity and multivariate stances of socio-technical environments, concept clarification and differentiation, finding common ground in language programming and transfer of terms in both ways from human to machine and vice versa and a specific methodology for evaluation which still appears to be in trial mode.

Waschull and Emmanouilidis (2023) proposed three dimensions of trustworthiness evaluation as a key quality feature for HCAI methodology exercise: technical, operational and social. The authors underlined for each dimension specific criteria such as accuracy, robustness, latency, reliability, security and scalability for the technical part; productivity, processing/manufacturing/configuration time, machine/process/product/routing flexibility, delivery reliability, produced and perceived quality and costs for the operational part; privacy, accountability, transparency, control, fairness and non-discrimination, inclusive character, variety, autonomy, ergonomics, social support, feedback, problem solving, work conditions, physical and mental demands, information needs for the social part (Waschull and Emmanouilidis 2023). Their study results also highlighted four major methodological steps which can guide an AI system design under the rule of ethics: 1. defining analysis units and case identification; 2. involvement of multidisciplinary teams and relevant stakeholders; 3. identification and validation of relevant performance units; 4. evaluation and feedback with focus on all previous steps.

From this point of view, every evaluation methodology for the exploration of HCAI is bound to be built on a stronger consideration for social variables, factors and criteria alongside technical objectives. Moreover, the new improved architecture of HCAI must shift the technic-centricity and

integrate it in social-technical overall systems which place human centricity at its core and not the other way around. Bingley et al. (2023) presented a social self-determination model for the AI system with positive or negative impact on the person inside a context, which combines elements from self-determination, social identity and self-categorization theories. The study presented a model in which AI systems can influence the well-being of an individual through self-determination and with a mediating role of individual or collective needs and identities (Bingley et al. 2023).

The Social Self Determination Model (SSDM) for AI environments opens the path for a better understanding and conceptualization of HCAI, ensuring new potentials for measure, methodology approaches, criteria evaluation and prediction for such a design. When working with intelligent systems, the individual at the workplace will likely exercise both an individual identity and a social identity depending on the personal needs, perceptions, expectations, personality cues, motivations, habits, values and previous experience. For example, accepting the partnership with a smart device can result in a resistance attitude and avoidance if the person will not perceive the collaboration in a personal development key, while ensuring security and privacy, objective evaluation and expected performance.

On the other hand, if the employee has a high level of trust towards AI, confidence in the fact that human boundaries will not be overshadowed by technical steps and witnesses in place clear regulations that function will more likely accept and engage in collaboration with an AI environment for task decision and process. Pogan and Popa (2020) showed in study that for the case of intelligent training platforms, the perception of usefulness, alongside level of experience, may shape employee attitudes and fears towards new technologies. More specific, an employee user with low experience in working with AI, who perceives the technology as threating, not fully explained or not human centred, will develop a negative perception and attitude towards it. Supplementary, any AI application development should follow personal freedoms, data protection and privacy (Cambon, 2017), ensuring human

centricity, in terms of ethical and moral principles and human well-being insurance.

As a general background, HCAI methodology is still in its prime, with often too many complex variables to be accounted for, too few articles and research endeavours that favour human dimensions inside intelligent systems and a lack of validated evaluation instruments and specific criteria for investigating such a sector. There is a great interest towards the future of HCAI research, knowing that the classical domains of implementation will surely expand, new comers adding to the industry, medical, communication, transportation and educational fields alongside new evaluation methods and measures of high validity destined for a proper scientific exploration and confirmation in this case.

Future directions for HCAI social approach and practical implications

Based on the actual research trends and results, it can be reaffirmed that future studies must take into consideration much more social dimensions and base socio-technic paradigms in their approach towards the creation, management, development and implementation of intelligent systems at work. One red line consists of testing more in-depth human variables in various experimental designs with regard to AI in specific and complex contexts. The moment for simple user-application design and interaction study has been surpassed by the need to understand human behaviour, personality, needs, expectations, emotions, perceptions and objectives in AI settings with focus on both actors but with a well-being status preference in the person's case.

Another future direction for the HCAI social approach solicits beyond the complex and in-depth human variable testing, a more rigorous concept and theory clarification with implications both in academic research and methodological premises. At the moment, too many heterogenous results, solitary one concept studies, the lack of a certain common ground and the dramatic change in AI evolution generates confusion and blurry outcomes. Probably with a clear guideline in place, specific theoretical explanations,

more replicated studies, validity and fidelity in testing, interconnected studies and various stakeholders collaboration, human – AI interaction can set in motion human centricity values and results.

HCAI social approach impacts the ethical and legal framework for future regulations and policies which regulate not only the creation, development and usage of AI but also the human principles and values beyond economical or simple technical demands. Already, a lot of countries are lining behind drawing such legal frames and directives which will adjust, orient, control and evaluate such systems with regard to people as central focal point inside the equation.

Among practical implications that outline the HCAI topic, one distinguishes through the numerous and multivariate character of applications in the Industry 4.0 with an increased awareness and attention given to the human user. Following human centricity in working with AI puts a spot on human values, integrating natural work environments for people, improving their development and life status, enhancing their competencies and appreciation in relation with an artificial intelligent actor that functions as a team member.

Another practical implication describes the use of HCAI for developing future smart systems at work that are self-sufficient and self-aware in order to support people functionality improvement, collaboration status, efficiency, interaction and well-being.

Last but not least, organizational development can benefit from human – AI interaction and collaboration in a social key, which ensures a hybrid work pattern for the future, with more advantages for company adaptation to dynamic markets, community participation, work-life balance, strong motivation and satisfaction level among employees. If reachable, such a stance can help generate a sense of self-fulfilment, security and trust for people in the work environment or context, opening the way for a new perception upon team work, decision making and risk taking in the organization.

Practical implications of HCAI reside in a new social perspective and methodology which place human values at the core of technology, investing ethical principles and human objectives inside intelligent systems.

Conclusions

HCAI paradigm has been investigated in the last years on a variety of topics, research areas and methodologies. Still the study exploration and international databases present too few examples on the matter. The shift from technology centred and traditional user variables towards new and in-depth human variables, focused more on perceptions, attitudes, cognition, personality and mentalities with a concrete purpose to place the person at the core of the technology has still a long journey to overtake. So far the study on AI has underlined from a social perspective the perception formation of employees and attitudes towards intelligent systems alongside team interaction in a mixed hybrid cooperation between man and machine.

Moreover, social sciences have explored various sectors of AI application and usage, from the industry domain, medical area and transportation to communications, education and entertainment, showing new and interesting results that need to be extended, validated and researched further on. Concept clarification and theory building are other facets that need to be addressed while HCAI imposes a high end complexity and dimension variety.

Future appointed implementation of HCAI will solicit more rigorous methodology with focus on common ground and regulations, specific guidelines, theoretical explanations, replicated studies, validity and fidelity testing, interdisciplinary character, stakeholders participation and ethical settings.

Social approach will have to shape new policies and legal framework for usage of AI in terms of human principles and values beyond simple economical or technical design objective and purpose. One of the major outcomes when dealing with AI in human centricity rules refers to developing

employee work and life quality, enhancing their skills, offering numerous possibilities for training, efficiency, respect towards diversity, transparency and equity.

Moreover, creating a sense of self-fulfilment, security and trust for people in a work context can become one of the main objectives of HCAI applications opening new paths for team work, problem solving, performance, motivation and satisfaction.

In this light of things, the social perspective and methodology can complete and enhance human centricity inside intelligent systems with respect to human values, ethical and moral grounds and an improved well-being.

References

Abdelaal, Abdelnasser. 2021. "Grand Research Challenges Facing Ethically Aligned Artificial Intelligence." In *Proceedings of the 27th American Conference on Information Systems (AMCIS)*, Virtual, 1-10. Montreal.

Amershi, Saleema, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, et al. 2019. "Guidelines for Human-AI Interaction." In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-13. Glasgow Scotland Uk: ACM. https://doi.org/10.1145/3290605.3300233.

Berente, Nicholas, Bin Gu, Jan Recker, and Radhika Santanam. 2021. "Managing Artificial Intelligence." *MIS Quaterly* 45(3):1-41. https://doi.org/10.25300/MISQ/2021/16274.

Bingley, William J., Caitlin Curtis, Steven Lockey, Alina Bialkowski, Nicole Gillespie, S. Alexander Haslam, Ryan K.L. Ko, Niklas Steffens, Janet Wiles, and Peter Worthy. 2023. "Where Is the Human in Human-Centered AI? Insights from Developer Priorities and User Experiences." *Computers in Human Behavior*, 141:1-8. https://doi.org/10.1016/j.chb.2022.107617.

Bingley, William J., S. Alexander Haslam, Niklas K. Steffens, Nicole Gillespie, Peter Worthy, Caitlin Curtis, Steven Lockey, Alina Bialkowski, Ryan K.L. Ko, and Janet Wiles. 2023. "Enlarging the Model of the Human at the Heart of Human-Centered AI: A Social Self-Determination Model

of AI System Impact." *New Ideas in Psychology*, 70: 1-12. https://doi.org/10.1016/j.newideapsych.2023.101025.

Boada, Júlia Pareto, Begoña Román Maestre, and Carme Torras Genís. 2021. "The Ethical Issues of Social Assistive Robotics: A Critical Literature Review." *Technology in Society,* 67:1-13. https://doi.org/10.1016/j.techsoc.2021.101726.

Braganza, Ashley, Weifeng Chen, Ana Canhoto, and Serap Sap. 2021. "Productive Employment and Decent Work: The Impact of AI Adoption on Psychological Contracts, Job Engagement and Employee Trust." *Journal of Business Research,* 131:485-494. https://doi.org/10.1016/j.jbusres.2020.08.018.

Cambon, Linda. 2017. "Health Smart Devices And Applications…Towards A New Model Of Prevention?." *European Journal of Public Health*, 27 (3):390-391. https://doi.org/10.1093/eurpub/ckx019.

Chowdhury, Soumyadeb, Prasanta Dey, Sian Joel-Edgar, Sudeshna Bhattacharya, Oscar Rodriguez-Espindola, Amelie Abadie, and Linh Truong. 2023. "Unlocking the Value of Artificial Intelligence in Human Resource Management through AI Capability Framework." *Human Resource Management Review,* 33 (1):1-21. https://doi.org/10.1016/j.hrmr.2022.100899.

Del Giudice, Manlio, Veronica Scuotto, Beatrice Orlando, and Mario Mustilli. 2023. "Toward the Human – Centered Approach. A Revised Model of Individual Acceptance of AI." *Human Resource Management Review,* 33(1):1-10. https://doi.org/10.1016/j.hrmr.2021.100856.

Dwivedi, Yogesh K., Nir Kshetri, Laurie Hughes, Emma Louise Slade, Anand Jeyaraj, Arpan Kumar Kar, Abdullah M. Baabdullah, et al. 2023. "Opinion Paper: 'So What If ChatGPT Wrote It?' Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy." *International Journal of Information Management,* 71: 1-63. https://doi.org/10.1016/j.ijinfomgt.2023.102642.

Flathmann, Christopher, Beau G. Schelble, Patrick J. Rosopa, Nathan J. McNeese, Rohit Mallick, and Kapil Chalil Madathil. 2023. "Examining the Impact of Varying Levels of AI Teammate Influence on Human-AI Teams." *International Journal of Human-Computer Studies*, 177:1-16. https://doi.org/10.1016/j.ijhcs.2023.103061.

Giermindl, Lisa Marie, Franz Strich, Oliver Christ, Ulrich Leicht-Deobald, and Abdullah Redzepi. 2022. "The Dark Sides of People Analytics: Reviewing the Perils for Organisations and Employees." *European*

Journal of Information Systems, 31(3):410-435. https://doi.org/10.1080/0960085X.2021.1927213.

Haleem, Abid, Mohd Javaid, and Ibrahim H. Khan. 2019. Current Status and Applications of Artificial Intelligence (AI) in Medical Field: An Overview. *Current Medicine Research and Practice*, 9(6):231-237. https://doi.org/10.1016/j.cmrp.2019.11.005.

Heyder, Teresa, Nina Passlack, and Oliver Posegga. 2023. "Ethical Management of Human-AI Interaction: Theory Development Review." *The Journal of Strategic Information Systems,* 32(3):1-50. https://doi.org/10.1016/j.jsis.2023.101772.

Hua, David, Neysa Petrina, Noel Young, Jin-Gun Cho, and Simon K. Poon. 2024. "Understanding the Factors Influencing Acceptability of AI in Medical Imaging Domains among Healthcare Professionals: A Scoping Review." *Artificial Intelligence in Medicine* 147:1-14. https://doi.org/10.1016/j.artmed.2023.102698.

Hu, Peng, Yaobin Lu, and Yeming Gong. 2021. "Dual humanness and trust in conversational AI: A person-centered approach." *Computers in Human Behavior*, 119:1-18. https://doi.org/10.1016/j.chb.2021.106727.

Islam, Gazi, and Michelle Greenwood. 2021. "Reconnecting to the Social in Business Ethics." *Journal of Business Ethics* 170 (1): 1-4. https://doi.org/10.1007/s10551-021-04775-7.

Kandul, Serhiy, Vincent Micheli, Juliane Beck, Thomas Burri, François Fleuret, Markus Kneer, and Markus Christen. 2023. "Human Control Redressed: Comparing AI and Human Predictability in a Real-Effort Task." *Computers in Human Behavior Reports,* 10:1-7. https://doi.org/10.1016/j.chbr.2023.100290.

Li, Jiahao, Yang Chu, and Jie Xu. 2023. "Impression Transference from AI to Human: The Impact of AI's Fairness on Interpersonal Perception in AI-Mediated Communication." *International Journal of Human-Computer Studies*, 179:1-11. https://doi.org/10.1016/j.ijhcs.2023.103119.

Li, Jian, and Jin-Song Huang. 2020. "Dimensions of Artificial Intelligence Anxiety Based on the Integrated Fear Acquisition Theory." *Technology in Society,* 63:1-10. https://doi.org/10.1016/j.techsoc.2020.101410.

Mirbabaie, Milad, Felix Brünker, Nicholas R. J. Möllmann Frick, and Stefan Stieglitz. 2022. "The Rise of Artificial Intelligence – Understanding the AI Identity Threat at the Workplace." *Electronic Markets* 32 (1): 73-99. https://doi.org/10.1007/s12525-021-00496-x.

Pogan, Livia and Radu-Ioan Popa. 2020. "Living In A Smart World. A Study Over Employees` Perceptions On The Usage Of Intelligent Devices." *Universitary Journal of Sociology*, 16(2):121-129.

http://www.sociologiecraiova.ro/revista/2020/12/revista-universitara-de-sociologie-no-2-2020/.

Restrepo Amariles, David, and Pablo Marcello Baquero. 2023. "Promises and Limits of Law for a Human-Centric Artificial Intelligence." *Computer Law & Security Review,* 48:1-10. https://doi.org/10.1016/j.clsr.2023.105795.

Sharma, Gagan Deep, Anshita Yadav, and Ritika Chopra. 2020. Artificial Intelligence and Effective Governance: A Review, Critique and Research Agenda. *Sustainable Futures*, 2:1-6. https://doi.org/10.1016/j.sftr.2019.100004.

Shneiderman, Ben. 2020a. "Human-Centered Artificial Intelligence: Three Fresh Ideas." *AIS Transactions on Human-Computer Interaction*, 12(3):109-124. https://doi.org/10.17705/1thci.00131.

Shneiderman, Ben. 2020b. "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):1-31. https://doi.org/10.1145/3419764.

Shneiderman, Ben. 2021. "Human-Centred AI." *Issues in Science and Technology*, 37(2):56-61.

Siau, Keng, and Weiyu Wang. 2020. "Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI." *Journal of Database Management,* 31(2):74–87. https://doi.org/10.4018/JDM.2020040105.

Te'eni, Dov, Michel Avital, Alan Hevner, Mareike Schoop, and David Schwartz. 2019. "It Takes Two to Tango: Choreographing the Interactions Between Human and Artificial Intelligence." In *Proceedings of the 27th European Conference on Information Systems: Information Systems for a Sharing Society (ECIS)*, Stockholm, Sweden.

Vishwarupe, Varad, Shrey Maheshwari, Aseem Deshmukh, Shweta Mhaisalkar, Prachi M. Joshi, and Nicole Mathias. 2022. "Bringing Humans at the Epicenter of Artificial Intelligence: A Confluence of AI, HCI and Human Centered Computing." *Procedia Computer Science*, 204:914-921. https://doi.org/10.1016/j.procs.2022.08.111.

Waschull, Sabine, and Christos Emmanouilidis. 2023. "Assessing Human-Centricity in AI Enabled Manufacturing Systems: A Socio-Technical Evaluation Methodology." *IFAC-PapersOnLine,* 56(2): 1791-1796. https://doi.org/10.1016/j.ifacol.2023.10.1891.

Yang, Stephen J.H., Hiroaki Ogata, Tatsunori Matsui, and Nian-Shing Chen. 2021. "Human-Centered Artificial Intelligence in Education: Seeing the Invisible through the Visible." *Computers and Education: Artificial Intelligence,* 2:1-5. https://doi.org/10.1016/j.caeai.2021.100008.

Zhu, Wenmin, Xiumin Fan, and Yanxin Zhang. 2019. "Applications and Research Trends of Digital Human Models in the Manufacturing Industry." *Virtual Reality & Intelligent Hardware*, 1(6):558-579. https://doi.org/10.1016/j.vrih.2019.09.005.