The Impact of Human Implication for AI-Supported Decisions Over Perception of Trust, Agency and Dignity

Camille Zinopoulos[1], Adam Fahmi[1], Sophie Boudreault[1], and Alexandre Marois[1,2]

[1]Université Laval, Québec, QC, G1V 0A6, Canada
[2]University of Central Lancashire, Preston, LAN, PR1 2HE, United Kingdom

ABSTRACT

Recent developments in artificial intelligence (AI), more specifically in generative AI, are disrupting our life. The integration of generative AI raises questions pertaining not only to the performance and accuracy of the AI system, but also to the boundaries of the role of both human and AI. This calls for a better understanding of the perception of human dignity over different uses of generative AI, but also for comprehending how said perception may interact with trust into the AI and sense of agency. The goal of the current study was to evaluate the perception of human dignity, trust and sense of agency among different uses of AI-supported decisions depending on the context of use and on the level of implication of the human decision maker. We presented participants a series of vignettes where generative AI systems were used to support decision making in five domains of use (health, business, humanities, arts, and technology) and four types of support (for decision support, communication, creativity, and research). The level of human implication regarding the decision was also manipulated across two conditions. Sense of agency, trust in the AI, perception of appropriateness for the AI to make a decision, as well as interpersonal justice and dehumanization level measures were collected for each vignette. Results outlined that sense of agency differed across conditions. Domain of use influenced sense of agency, trust in the AI, decision appropriateness and dehumanization perceptions, with differences emerging mostly for health-related vignettes. The type of support also impacted trust and decision appropriateness, with more positive perceptions for vignettes discussing creativity use cases. Overall, our study sheds light on the perception of the general population over different types of AI use and how components such as perception of agency, trust and dignity may vary depending on the nature of the use.

Keywords: Generative AI, Human dignity, Trust in AI, Sense of agency, Questionnaire

INTRODUCTION

Within the last decade, we have witnessed important technological developments. More specifically, information technologies have increasingly improved, both regarding their computing capabilities and the diversity of tasks they can now carry on. Among these developments, artificial intelligence (AI) has become omnipresent within our society, becoming a

subject of discussion. AI technologies encompass many different tools, but generative AI represents one of the main technologies that has potential for changing the role of humans in many different situations. Generative AI is a technology that can be used to automatically and rapidly create content that is typically produced by humans (Cao et al., 2023). Generative AI is a key part of many tools including chatbots such as ChatGPT, developed by OpenAI, which can understand human language and produce responses in coherence with the subject discussed (Feuerriegel et al., 2023). The efficiency of generative AI-driven technologies and the quality of the content they create with respect to what can be produced by humans opens for many opportunities to alleviate human tasks across a variety of domains (see, e.g., Gupta et al., 2024, for a review). Yet, the integration of generative AI raises many questions and concerns. Not only is it legitimate to question the technical advances of these technologies (e.g. regarding the capabilities that generative AI can achieve), but several considerations central to the humans need to be investigated (Dwivedi et al., 2023). This includes better understanding and clarifying where lies the frontier between human and machine roles, as well as reflecting upon respecting and upholding human dignity (Kreps & Rowe, 2021). From a sociotechnical point of view (Storey et al., 2020), technological developments must consider societal concerns. Consequently, research focusing on the role of humans in relation to such disruptive technologies must progress concomitantly to technical advances. The general aim of this study is to assess the perception of human dignity over different uses of generative AI.

Human Dignity and Human-Autonomy Teaming Considerations

Human dignity refers to a universal form of respect for basic human rights. This principle is central to the Universal Declaration of Human Rights as outlined in its first article: "All human beings are born free and equal in dignity and rights" (United Nations, 1948). From a philosophical perspective, human dignity is a paramount component to human experience. According to De Koninck (2005), respect for human dignity is closely related to respect for intelligence, will and freedom. From a psychological point of view, human dignity—or lack thereof—is related to experiences of shame and humiliation (Hojman & Miranda, 2018).

Human dignity has been scarcely studied in the context of human-autonomy teaming (HAT), which represents an appropriate framework for understanding how human and AI can interact—cooperate even—which each other. Nevertheless, the recent democratization of AT technologies has pushed forward reflections on how AI-driven systems might affect human dignity in different contexts (Sathl et al., 2023). Formosa et al. (2022) assessed whether the impersonal nature of AI could give rise to concerns regarding human dignity respect in a medical context. As such, they asked participants to read approximately 20 clinical vignettes that presented scenarios in which diagnostic services or medical resource allocation were offered by either a human or an AI. For each vignette, participants reported their opinion on various components of human dignity including, for

instance, perceived justice, satisfaction and dehumanization of the decision, as well as assessing the level of trust and appropriateness of the entity responsible for the decision. The results showed a systematic positive bias towards decisions made by humans and a feeling of dehumanization towards AI decisions. Trust in the AI decision was also higher when the decision entailed a positive outcome and positively correlated with a better perception of justice (see also Bankins et al., 2022). These results are in line with other studies outlining that AI is thought to produce decisions while neglecting human considerations (Binns et al., 2018)—even relaying them as mere depersonalized data—and that it is inappropriate for making moral decisions (Formosa & Ryan, 2021).

As outlined by Formosa et al. (2022; see also Bankins et al., 2022), it seems that dignity is also closely related to trust. Trust is an essential component of many different HAT models as it can highly influence the nature of the collaboration between both human and machine components (de Visser et al., 2018). Yet, HAT principles such as trust are often taken for granted and not always considered when it comes to studying how particular technologies might be perceived. This is also the case for the sense of agency, that is the perception of control of a user over a given tool used in a particular situation. Agency is paramount for establishing one's engagement for using a tool as well as for understanding how one perceives the locus of control of their role with respect to the technology they rely on (Wen & Imamizu, 2022). Again, this speaks to the definition of human dignity given that, for both of these concepts, the place that humans take and the feeling of their freedom to act and decide on a given action are central. Besides, HAT experience can vary widely depending on which component makes the final decision (i.e. either human or AI). It is essential to better understand how trust, agency and human dignity considerations can be interrelated with respect to the use of generative AI technologies. This would allow better framing and anticipating the integration of such tools in society, and to eventually understand how they can be accepted for content creation under different use cases.

The Present Study

The goal of the current study was to evaluate the perception of human dignity, trust and sense of agency among different uses of AI-supported decisions depending on the context of use and on the level of implication of the human decision maker. To reach this goal, we presented participants a series of vignettes where generative AI systems were used to support decision making in five domains of use (health, business, humanities, arts, and technology). For each of these domains, different vignettes were presented where the level of human implication regarding the decision differed (i.e. the final decision being either made by the human with advice from an AI or almost totally made by the AI decision support).

METHOD

Participants

One hundred and three participants ($M_{age} = 36.88$) recruited from social networks or from Université Laval, Université du Québec à Chicoutimi or Université de Sherbrooke took part in this study in exchange for a chance to win a 50$ gift card. All participants reported being able to complete a French-written survey.

Apparatus and Material

The survey was presented on LimeSurvey. Participants were shown 20 vignettes and were asked to answer a set of items regarding each of these vignettes. The vignettes took the form of brief hypothetical scenarios that involved discussing with an AI-driven conversational tool that could provide answers and support for specific purposes. Among the 20 vignettes, different domains of use (5) and types of AI-driven support (4) were described. The five domains of use went as follows: a) health sciences; b) business; c) social sciences and humanities; d) arts and entertainment; and e) technology and infrastructure. The four types of support included: a) assistance and decision support; b) communication and interaction; c) creativity; and d) research and analysis. This thus resulted in 20 combinations of domains of use and types of support (i.e. $5 \times 4 = 20$). Among the 103 participants, 58 were shown vignettes where human implication was higher whereas 45 subjects were shown vignettes characterized by a lower role of the human in the decision making. Below is an example of a vignette with higher human involvement for the business domain and support for creativity (High human involvement condition):

> You finally realize your childhood dream by launching your own small artisan pastry business. Aware of the importance of a distinctive visual identity, you decide to create a unique logo that reflects your passion for pastry. Lacking the necessary design skills, you decide to use an AI. Based on your preferences, it generates several logos, allowing you to choose the one that best embodies the spirit of your future company. To check the quality of your logo, you ask a graphic designer friend for advice.

Below is the same example for the business domain and support for creativity but with higher involvement of the AI tool (Low human involvement condition):

> You finally realize your childhood dream by launching your own small artisan pastry business. Aware of the importance of a distinctive visual identity, you decide to create a unique logo that reflects your passion for pastry. Lacking the necessary design skills, you decide to use an AI. Based on your preferences, it generates a logo that you use instantly to represent the spirit of your company.

Following each vignette, a set of five questionnaires was presented. (1) The *Sense of Agency Scale* (Tapal et al., 2017) was used to measure sense of agency with respect to the decision being made in the scenario. This included 10 items regarding the sense of control of the decision maker. Each item was answered on a 1-to-7 Likert scale (from highly disagree to highly agree). (2) The *Trust in Automation Scale* (Körber, 2018) is a two-item scale aiming at measuring trust in a machine. The items were adjusted to refer to an AI (e.g., "I'd trust the AI described in the vignette"). A five-point Likert scale (from highly disagree to highly agree) was used to answer the questionnaire. (3) The *Decision-Maker Role Appropriateness* (Bankins et al., 2022) was used to assess the opinion of participants regarding the appropriateness of the AI to decide in the vignettes depicted. This one-item measure (i.e. "In the scenario, to what extent is it appropriate for the AI described to make a decision") was answered on a seven-point Likert scale (from very inappropriate to very appropriate). (4) The *Interpersonal Justice* questionnaire of Bies & Moag (1986) aimed at measuring how participants could feel during the hypothetical interaction with the AI described in the vignettes. It consisted of four items related to respect, dignity and politeness, answered on a five-point Likert scale (from a low level of respect to a high level). (5) Finally, the *Human Nature* items of the *Dehumanization* questionnaire (Bastian & Haslam, 2011) was modified to measure how participants perceived the way the AI could treat them with respect of their human nature. This included five questions to be answered on a five-point scale (from a low-level respect to a high level).

Procedure

Participants received the link for the survey after having raised their interest to partake in the study. They were randomly assigned a condition (either High human involvement or Low human involvement condition) and were sent the survey accordingly. After reading the consent form and a brief description of the survey, they were presented the 20 vignettes, each accompanied by the five questionnaires. Vignettes were presented in a random order across participants. After having answered all the vignettes, they were asked sociodemographic questions and were thanked for their participation.

Analysis

The impacts of the domain of use and type of support were both assessed while also considering the effect of human involvement. To do so, measures on the questionnaires were averaged to represent the mean of all vignettes concerning either common domain of use or type of support. Then, a series of mixed analyses of variance (ANOVAs) with the between-subjects factor Human involvement (High human involvement vs. Low human involvement) and either the Domain of use (health sciences; business; social sciences and humanities; arts and entertainment; and technology and infrastructure) or Type of support (assistance and decision support; communication and interaction; creativity; research and analysis) was conducted. These mixed

ANOVAs were performed on the following variables: Sense of agency mean score, Trust mean score, Decision-Maker Appropriateness mean score, Interpersonal justice mean score, and Dehumanization mean score. To look for potential relationships across variables, Pearson correlation analyses were also performed between each of these variables but collapsed across all domains and use and types of support. These correlational analyses were performed within each group.

RESULTS

Table 1 depicts the mean values and standard deviations for each of the measure collected, averaged across domains of use for both low and high human involvement conditions. As shown in the table below, different domains of use seemed to induce different perceptions over the measures collected. To test these differences, mixed ANOVAs were carried out for each measure. For the Sense of agency measure, significant main effects were found for the Domain of use, $F(4, 196) = 2.46$, $p = .047$, $\eta^2_p = .05$, and for the Condition, $F(1, 49) = 7.88$, $p = .007$, $\eta^2_p = .14$. The interaction failed to reach significance ($F = 1.08$, $p = .367$). Sense of agency was higher for vignettes characterized by higher human involvement ($p = .007$) and higher in the Humanities vignettes as opposed to the Health vignettes ($p = .029$). The mixed ANOVA performed over the measures of trust in the AI described raised a significant effect of Domain of use, $F(4, 196) = 28.27$, $p < .001$, $\eta^2_p = .37$, and a significant Domain of use × Condition interaction, $F(4, 196) = 4.27$, $p = .005$, $\eta^2_p = .08$. The effect of Condition did not reach significance ($F = .05$, $p = .830$). Decomposition of the interaction showed that trust in health-related decisions was significantly lower than all other domains across both conditions ($ps < .042$). For the Low human involvement condition, trust in the AI was lower in the business vignettes compared with humanities and technology vignettes ($ps < .028$). Analysis of the Decision-maker role appropriateness raised a main effect of Domain of use, $F(4, 196) = 27.81$, $p < .001$, $\eta^2_p = .36$, but no main effect of Condition nor of two-way interaction (with $F = 1.05$, $p = .312$, and $F = 1.66$, $p = .175$, respectively). Posthoc analyses of the impact of the domain of use outlined that the appropriateness of the AI to decide was systematically lower for health vignettes ($ps < .001$). Other domains did not differ. Analysis of the Interpersonal justice showed no effect of Domain of use, Condition, nor any two-way interaction ($Fs < 1.81$, $ps > .152$). Finally, the mixed ANOVA performed on the Dehumanization measure raised a significant main effect of Domain of use, $F(4, 192) = 3.17$, $p = .028$, $\eta^2_p = .06$, and no Condition or interaction effect ($Fs < 0.95$, $ps > .422$). Perceptions of respect was lower for the health-related vignettes as opposed to the business-related vignettes ($p = .026$). Other comparisons did not reach significance.

Table 1. Mean values and standard deviations (in parentheses) for the measures across domains of use and human involvement conditions.

Measure	Domain of Use				
	Health	Business	Humanities	Arts	Technology
Low human involvement					
SoA	4.58 (1.09)	4.31 (1.08)	4.75 (1.03)	4.65 (1.23)	4.67 (1.01)
Trust	2.60 (1.03)	3.13 (0.90)	3.40 (0.92)	3.46 (0.95)	3.59 (0.90)
DM role	2.83 (1.56)	3.96 (1.45)	4.28 (1.66)	4.44 (1.55)	4.64 (1.56)
Justice	4.08 (0.83)	4.11 (0.93)	4.03 (0.92)	4.11 (0.87)	3.97 (1.06)
Dehum.	2.84 (1.10)	3.04 (1.09)	3.07 (1.19)	2.99 (1.17)	3.01 (1.18)
High human involvement					
SoA	5.40 (0.92)	5.44 (0.86)	5.16 (1.04)	5.14 (1.21)	5.34 (1.01)
Trust	2.87 (1.00)	3.67 (0.88)	3.36 (0.81)	3.71 (1.04)	3.30 (0.84)
DM role	3.26 (1.50)	4.64 (1.47)	4.59 (1.55)	4.84 (1.58)	4.57 (1.46)
Justice	4.09 (1.04)	4.24 (1.09)	4.12 (0.92)	4.13 (1.00)	4.11 (1.15)
Dehum.	3.14 (1.18)	3.24 (1.33)	3.20 (1.22)	3.19 (1.28)	3.18 (1.27)

Note. SoA: Sense of agency; DM role: Decision-maker appropriateness role; Dehum.: Dehumanization. SoA and DM role are measured on a 7-point scale whereas other measures are on a 5-point scale.

Table 2 presents the mean values and standard deviations for the different measures collected across the different types of support presented in the vignettes as a function of the human involvement conditions. The Sense of agency only differed across conditions as supported by the significant effect of Condition, $F(1, 54) = 12.70$, $p < .001$, $\eta^2_p = .19$, and the absence of effect of Type of support ($F = 0.18$, $p = .909$) and of interaction ($F = 2.41$, $p = .069$). Perception of agency was higher for the High-human involvement condition. Trust in the AI described in the vignettes varied across Types of support with a significant effect of this factor, $F(3, 162) = 6.36$, $p = .001$, $\eta^2_p = .11$. However, measures of trust were not impacted by the Condition and no interaction emerged ($Fs < 1.51$, $ps > .215$). Further analysis showed that trust was superior for all the Creativity vignettes as opposed to all other types of support ($ps < .016$). Measures of trust across other types of support failed to differ. Appropriateness of the decision being made by an AI also varied across types of support as supported by a main effect of Type of support, $F(3, 162) = 6.42$, $p < .001$, $\eta^2_p = .11$. Multiple comparisons showed that participants considered the implication of the AI as being more appropriate for creativity-related decisions as opposed to all other types of decisions ($ps < .011$). No other comparison reached significance, and no main effect of Condition or interaction arose ($Fs < 1.58$, $ps > .202$). Finally, measures of Interpersonal justice and Dehumanization remained similar across the Types of support and Conditions with an absence of main effects and of interaction for both these measures ($Fs < 1.93$, $ps > .139$).

Table 2. Mean values and standard deviations (in parentheses) for the measures across types of support and human involvement conditions.

Measure	Type of Support			
	Decision	Communication	Creativity	Research
Low human involvement				
SoA	4.73 (1.09)	4.52 (1.16)	4.57 (1.12)	4.58 (1.04)
Trust	3.15 (0.95)	3.20 (0.82)	3.35 (0.91)	3.22 (0.92)
DM role	3.88 (1.44)	3.98 (1.52)	4.38 (1.61)	3.95 (1.54)
Justice	4.13 (0.92)	4.14 (0.88)	3.95 (1.01)	4.04 (0.88)
Dehum.	3.03 (1.15)	2.88 (1.14)	2.87 (1.18)	3.00 (1.11)
High human involvement				
SoA	5.43 (1.04)	5.23 (0.92)	5.15 (0.95)	5.34 (1.05)
Trust	3.35 (0.97)	3.21 (0.96)	3.66 (0.84)	3.31 (1.03)
DM role	4.22 (1.61)	4.09 (1.60)	4.45 (1.70)	4.14 (1.77)
Justice	4.21 (1.00)	4.10 (0.95)	4.18 (1.07)	4.17 (1.05)
Dehum.	3.14 (1.30)	3.21 (1.18)	3.38 (1.24)	3.21 (1.25)

Note. SoA: Sense of agency; DM role: Decision-maker appropriateness role; Dehum.: Dehumanization. SoA and DM role are measured on a 7-point scale whereas other measures are on a 5-point scale.

To push further our comprehension of the relationship between all the different measures collected, we performed a set of Pearson correlation analyses between Sense of agency, Trust, Decision-maker role appropriateness, Interpersonal justice and Dehumanization within each condition. For the Low-human involvement condition, we found a significant positive relationship between Trust and Sense of agency ($r = .56$, $p < .001$), Trust and Decision-maker role appropriateness ($r = .79$, $p < .001$), and Trust and Interpersonal justice ($r = .33$, $p = .014$). Sense of agency was also positively correlated to the Decision-maker role appropriateness ($r = .57$, $p < .001$) and negatively correlated with the Dehumanization level ($r = -.31$, $p = .021$). Within the High-human involvement condition, Trust in the AI was also positively correlated with Sense of agency ($r = .33$, $p = .027$), Decision-maker role appropriateness ($r = .72$, $p < .001$), and Interpersonal justice ($r = .58$, $p = .014$). Sense of agency was also negatively correlated with the Dehumanization level ($r = -.36$, $p = .020$). However, contrary to the Low-human involvement condition, Interpersonal justice was positively associated with Sense of agency ($r = .42$, $p = .006$), and with Decision-maker role appropriateness ($r = .41$, $p = .008$).

DISCUSSION

The goal of this study was to provide a first assessment on how different uses of generative AI were perceived by the public. More precisely, participants were shown vignettes presenting various examples of domains of use, types of support, and degrees of human involvement in the final decision being made and were asked about their perception on a series of HAT and human dignity measures.

Vignettes characterized by higher human implication were related to higher levels of agency over the AI use described. Healthcare-related vignettes generally induced more negative perceptions that the other domains of use; they induced lower sense of agency as compared with social sciences and humanities vignettes, lower trust in the AI than all the other domains, and reduced appropriateness for the involvement of the AI in the decision compared with other domains. Perception of dehumanization was higher for heath vignettes, compared with business vignettes. Among the low-human involvement vignettes (i.e. with more decision power ascribed to the AI), trust was also lower for the business vignettes compared with humanities and technology vignettes. As for the type of support, it mostly impacted trust and the appropriateness of the involvement of the AI in the decision. For both measures, support for creativity tasks was more trusted and considered more appropriate than other types of support. Overall, these results show that generative AI is perceived differently depending on the nature of the task it is used for and the context in which it is used.

Correlation analyses also outlined how measures of trust, agency and decision-maker appropriateness as well as measures related to human dignity were related to each other with some differences depending on the level of human implication. For decisions being mostly made by the AI (low human involvement), the sense of agency and appropriateness level for the AI to decide were related, which was not the case for vignettes characterized by higher human involvement. For decisions being mostly made by the human (high human involvement), perceptions of justice were related to the appropriateness of the AI involvement and sense of agency. In both conditions, perceptions of dehumanization were negatively related to sense of agency. These results highlight that the relationship between all these measures of HAT and perceptions of human dignity is impacted by the extent to which human vs. AI roles are more or less potent.

A brief look at the different scales of each of the measures collected highlights that perceptions were seen as more positive or less negative, depending on the domain of use. For instance, in the case of health vignettes, the mean values for the trust, decision maker role appropriateness and dehumanization measures were under the midscale value, suggesting that using generative AI for supporting human decision in these situations was considered negative. Other domains were seen in a relatively more positive way. Sense of agency, although it was significantly lower in the low-human involvement condition, was higher than the midscale value for both conditions, meaning that participants still considered they possessed control over the situation in all the vignettes presented.

These results are consistent with those outlined in previous research conducted on the use of AI in certain domains. In Formosa et al. (2022), participants viewed AI-driven decisions in medicine as being untrustworthy and dehumanizing. In Bankins et al. (2022), similar conclusions were reached in a resource management context. The more negative views reported for health-related vignettes, and in some cases for business vignettes, are coherent with these results. The different correlations found across HAT and dignity measures shed new light over the perception of generative AI from the point

of view of a user. Though it has scarcely been studied, the relationships found are still coherent with Formosa et al. (2022) who showed a positive correlation between perceptions of justice and trust. The association between sense of agency and dignity, which has not necessarily been studied before, is also conceptually in line with how sense of agency is defined. As outlined by Wen and Imamizu (2022), agency is closely related to the view that one has over their own control and will over a situation and freedom to act, two elements that are essential to respect human dignity (De Koninck, 2005).

CONCLUSION

In conclusion, our study provides a first look at how different uses of generative AI are perceived by the general public with regard to different HAT and dignity principles. Results showed that health uses were perceived in a relatively different way than other uses. Creativity uses were also perceived as being more positive. Different relationships were found across the measures collected. These results represent a first towards better understanding how human experience changes when generative AI is involved in the decision made. They can help refining our view and integration of generative AI across different domains and contexts to ensure that users will feel considered, respected, and that they will actually trust the type of decision made. In light of the different relationships found, the next steps will involve to better understand how the components measured can be related to actual uses and intent of uses of the technologies described in order to anticipate, predict even, how certain uses of generative AI may be actually trusted and applied across different domains.

ACKNOWLEDGMENT

REFERENCES

Bankins, S., Formosa, P., Griep, Y. and Richards, D. (2022) "AI decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in a human resource management context", *Information Systems Frontiers*, *24*, 857–875. https://doi.org/10.1007/s10796-021-10223-8

Bastian, B. and Haslam, N. (2011) "Experiencing dehumanization: Cognitive and emotional effects of everyday dehumanization", *Basic and Applied Social Psychology*, *33*, 295–303. https://doi.org/10.1080/01973533.2011.614132

Bies, R. J. and Moag, J. F. (1986). "Interactional justice: Communication criteria of fairness", In: R. J. Lewicki, B. H. Sheppard, and M. H. Bazerman (Eds)., *Research on Negotiations in Organizations*, *1*, 43–55. JAI Press.

Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J. and Shadbolt, N. (2018) "'It's reducing a human being to a percentage'; Perceptions of justice in algorithmic decisions", *CHI'18 Conference on Human Factors and Computer Systems*, 377. https://doi.org/10.1145/3173574.3173951

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S. and Sun, L. (2023) "A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT", arXiv:2303.04226. https://doi.org/10.48550/arXiv.2303.04226

De Koninck, T. (2005) "Archéologie de la nation de dignité humaine" In: T. De Koninck and G. Larochelle (Eds.) , *La dignité humaine. Philosophie, droit, politique, économie, médecine* (pp.13–50). Presses universitaires de France.

de Visser, E., Pak, R. and Shaw, T. H. (2018) "From 'automation' to 'autonomy': the importance of trust repair in human-machine interaction", *Ergonomics*, *61*, 1409–1427. https://doi.org/10.1080/00140139.2018.1457725

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghvan, V., Ahuja, M., … Wright, R. (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*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Feuerriegel, S., Hartmann, J., Janiesch, C., and Zachech, P. (2023) "Generative AI", *Business & Information Systems Engineering,* *66*, 111–126. https://doi.org/10.1007/s12599-023-00834-7

Formosa, P., Rogers, W., Griep, Y., Bankins, S. and Richards, D. (2022) "Medical AI and human dignity: Contrasting perceptions of human and artificially intelligent (AI) decision making in diagnostic and medical resource allocation contexts", *Computers in Human Behavior*, *133*, 107296. https://doi.org/10.1016/j.chb.2022.107296

Formosa, P. and Ryan, M. (2021) "Making moral machines: why we need artificial moral agents", *AI & Society*, *36*, 839–851. https://doi.org/10.1007/s00146-020-01089-6

Gupta, P., Ding, B., Guan, C. and Ding, D. (2024) "Generative AI: A systematic review using topic modelling techniques", *Data and Information Management*, *8*, 100066. https://doi.org/10.1016/j.dim.2024.100066

Hojman, D. A. and Miranda, A. (2018) "Agency, human dignity, and subjective well-being", *World Development*, *101*, 1–15. https://doi.org/10.1016/j.worlddev.2017.07.029

Körber, M. (2019) "Theoretical considerations and development of a questionnaire to measure trust in automation", In: S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander and Y. Fujita (Eds.), *Proceedings of the 20th Congress of the International Ergonomics Association. IEA 2018. Advances in Intelligent Systems and Computing*, *823*. Springer. https://doi.org/10.1007/978-3-319-96074-6_2

Kreps, D. and Rowe, F. (2021) "Free will in philosophy and in IS research: where do we stand and where do we want to be?", *ICIS 2021 Proceedings*, *5*.

Sathl, B. C., Schroeder, D. and Rodrigues, R. (2023) "*Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges*". Springer.

Storey, M.-A., Ernst, N. A., Williams, C. and Kalliamvakou, E. (2020) "The who, what, how of software engineering research: a socio-technical framework", *Empirical Software Engineering*, *25*, 4097-4129. https://doi.org/10.1007/s10664-020-09858-z

Tapal, A., Oren, E., Dar, R. and Eitam, B. (2017) "The sense of agency: A measure of consciously perceived control over one's mind, body, and the immediate environment", *Frontiers in Psychology*, *8*, 1552. https://doi.org/10.3389/fpsyg.2017.01552

United Nations (1948) "Universal declaration of human rights", United Nations.

Wen, W. and Imamizu, H. (2022) "The sense of agency in perception, behaviour and human-machine interactions", *Nature Reviews Psychology*, *1*, 211–222. https://doi.org/10.1038/s44159-022-00030-6