Journal of Strategic Information Systems 32 (2023) 101772

Contents lists available at ScienceDirect

Journal of Strategic Information Systems

journal homepage: www.elsevier.com/locate/jsis

Ethical management of human-AI interaction: Theory development review

Teresa Heyder, Nina Passlack, Oliver Posegga *

Information Systems and Social Networks, University of Bamberg, An der Weberei 5, 96047 Bamberg, Germany

A R T I C L E I N F O

Keywords:
Artificial intelligence
Ethics
Human-AI interaction
Theoretical review
Sociomateriality

A B S T R A C T

AI-based technologies have changed the nature of the symbiosis between humans and AI, and so strategic management of human-AI interaction in organizations requires deeper ethical considerations. Aligning AI with human values requires a systematic understanding of the ethical management of human-AI interaction. We conduct a theoretical review, from a sociotechnical perspective, and analyze ethical management of human-AI interaction through the lens of sociomateriality. Our systematic approach helps explain and clarify the interdependencies between two ethical perspectives – duty and virtue ethics – in sociotechnical systems. We also provide a theoretical framework that leads to seven avenues for future research.

Introduction

The most recent generations of technologies based on Artificial Intelligence (AI), characterized by the ability "to perform cognitive functions that we associate with human minds" (Rai et al., 2019, p. iii), change the nature of the symbiosis between humans and technology: AI increases "human task performance with higher extents of interactivity and intelligence" than previous technologies (Maedche et al., 2019, p. 536). AI differs from previous technologies in its ability to act (semi-)autonomously (Maedche et al., 2019; Rieder et al., 2020; Scherer, 2016). For example, recent developments in generative AI, such as the growing maturity of Large Language Models (LLMs), increase the relevance and acceptance of AI-based technologies in organizations (Dwivedi et al., 2023; Markus and Rowe, 2023). With their growing capabilities, maturity, and pervasiveness, such technologies have the potential to reshape traditional forms of work, giving rise to "human–machine hybrid work" (Mollick, 2022, p. 5). AI-based technologies are "no longer always subordinate to the human agent, [as they] can now assume responsibility for tasks" (Baird and Maruping, 2021, p. 315). When AI-based systems act (semi-) autonomously (Rieder et al., 2020), autonomy shifts from humans to technical systems. This influences the collaboration of humans and technology (Baird and Maruping, 2021), raising ethical challenges such as distributive justice concerns, discrimination and exclusion, and transparency issues through the perceived powerlessness of humans as well as AI opacity (e.g., Abdelaal, 2021, Boada et al., 2021, Dwivedi et al., 2023, Giermindl et al., 2022).

To avoid unintended negative consequences of the ethical challenges that result from this changing nature of human-AI interaction, we must rethink how we manage human-AI interaction in organizations (Rai et al., 2019). To promote positives outcomes such as well-being, organizations must consider ethics and morals when making decisions about the implementation and use of the latest technologies based on AI to realize their strategic goals (Berente et al., 2021; Dwivedi et al., 2023; Marabelli et al., 2021). Considering

* Corresponding author.
E-mail addresses: teresa.heyder@uni-bamberg.de (T. Heyder), nina.passlack@uni-bamberg.de (N. Passlack), oliver.posegga@uni-bamberg.de (O. Posegga).

https://doi.org/10.1016/j.jsis.2023.101772

Available online 3 August 2023

ethics in strategic management requires two perspectives. On the one hand, managing human-AI interaction needs to enforce *ethics by duty* by establishing guidelines and developing principles that prescribe correct ethical behavior as an obligation for employees and AI (Alsheibani et al., 2020; Han et al., 2020; Rakowski et al., 2021). On the other hand, strategic management needs to be aware of the workforce's *virtue ethics* based on the inherent moral values that guide the intrinsic motivation of individuals to behave ethically when implementing, designing, using, and deciding upon AI (Flathmann et al., 2021). Hence, to foster ethical management of human-AI interaction, it is important to combine both ethical perspectives (Eitel-Porter, 2021; Siau and Wang, 2020).

Although recently there has been valuable research on ethics in the context of AI, there appear to be four gaps in research on the ethical management of human-AI interaction. First, the discourse lacks a general theory that explains the complexity of ethics and AI (Mirbabaie et al., 2022), that is, combining the technical discourse with ethical philosophy (Berente et al., 2021). Second, research has provided limited approaches to understanding contemporary ethical challenges at the crossroads of society, business, and technology (Islam and Greenwood, 2021). Third, the current discourse on ethics in the context of managing human-AI interaction is restricted mainly to the formulation of "principles, rules, guidelines, policies, and regulations for AI … to exhibit ethical behavior" (Siau and Wang, 2020, p. 74), representing only one of the perspectives (*duty ethics*) required to integrate ethics in strategic management. Finally, Information Systems (IS) research and research in closely related fields often adopt a technical stance and gravitate toward studying and designing technologies, including AI methods and algorithms, as well as their characteristics, such as fairness and explainability. They often fail to consider explicitly human actors and how their behavior might affect human-AI interaction, which is also required to understand and resolve ethical issues in workplace management of human-AI interaction (Amershi et al., 2019; Berente et al., 2021; Te'eni et al., 2019). In addressing those four gaps, we strive for an interdisciplinary, systematic, and holistic approach that addresses the following research question:

How can we understand and explain ethical management of human-AI interaction?

In order to do justice to the complexity of the phenomenon under research and to ensure a human-centered perspective, there is a need to reflect upon the management of human-AI interaction in both directions: managing AI interacting with humans (incorporating the *duty ethics* perspective of integrating ethical principles in the design of AI-based technologies); and managing humans interacting with AI (incorporating the *duty ethics* perspective of establishing ethical principles to guide their behavior when working with AI as well as the *virtue ethics* perspective aiming to align human values with AI). This leads to the two subquestions (SQs): *How can we understand and explain ethical management of AI interacting with humans?* (SQ1), and *How can we understand and explain ethical management of humans interacting with AI?* (SQ2).

To achieve our research goal, we conduct a *theoretical review*. This type of literature review aims at building an explanation by employing a systematic and structured approach to transforming constructs identified in previous work, spread across several research streams, into a conceptual framework that allows for examining its components and interdependencies (Paré et al., 2015). In terms of theory development, we adopt the *extend strategy* of "Next-Generation Theorizing" of Burton-Jones et al. (2021, p. 303) by using well-established theories to explain human-AI interaction as a "new phenomen[on] emerging in a changing world" (Burton-Jones et al., 2021, p. 303). To do so, our work builds on the conceptualization by Leonardi (2012), which summarizes the discourse on sociomateriality in a conceptualization of sociotechnical systems. The sociotechnical perspective allows us to reflect upon the current progress of AI-based technologies, as such technologies interfere with interpersonal relations and, thus, represent "sociotechnical systems par excellence" (Greene et al., forthcoming, p. 2). Sociotechnical system theory represents how interdependencies between humans, technologies, and other objects (such as guidelines and management practices) being arranged in organizations lead inter alia to ethical or non-ethical patterns of actions (Cecez-Kecmanovic et al., 2014; Contractor et al., 2011). We extend Leonardi's (2012) conceptualization through state-of-the-art literature on managing human-AI interaction in organizations that we identified in our systematic literature review. To provide a more nuanced analysis of this literature, we use the notions of sociomateriality (*inseparability*, *relationality*, *performativity*, *materiality,* and *practices*) as they "individually and collectively, offer a distinctive and coherent account of the relationships between" all parts (human, technologies, and other objects) being arranged in organizations (Jones, 2014, p. 895). These notions enable us to improve our understanding of ethical interdependencies arising through the "entanglement" of humans and technologies in organizations influenced by management practices and guidelines (Cecez-Kecmanovic et al., 2014). We thus go beyond synthesizing the current literature, aiming to deliver a comprehensive *theoretical review* to increase knowledge on the ethical management of human-AI interaction involving both ethical perspectives (*duty* and *virtue*).

We expect to contribute to research in three areas. First, we extend knowledge on the ethical management of human-AI interaction systematically, not only by consolidating existing research insights but also by explaining as sociotechnical systems the ethical interdependencies between humans, technologies, and other objects in organizations. Second, we respond to the call of Berente et al. (2021) by applying the sociotechnical perspective and showing that it is not only a product of theorizing (outcome) but also can guide ethical considerations in managing human-AI interaction. We agree with Greene et al. (forthcoming), that the IS field is responsible for and "uniquely qualified to evaluate the scientific, social, and technical aspects" of the newest technologies being based on AI and using unsupervised learning techniques (p. 2). Thus, we provide a systematic approach for analyzing the ethical management of human-AI interaction through the lens of sociomateriality. Third, we combine the technical discourse with ethical philosophy to account for both the *duty ethics* and *virtue ethics* perspectives, and thus, following the calls of Berente et al. (2021) and Mirbabaie et al. (2022), for a general theory that accounts for the complexity of ethics and AI. Thereby, we combine older and new theories, not only from IS research but also from interdisciplinary domains, thus showing how "Next-Generation Theorizing" may be conducted in the context of a systematic *theoretical review* (Paré et al., 2015; Templier and Paré, 2018). Furthermore, we heed the calls of Vial (2019) and Alsheibani et al. (2020) to enhance our understanding of ethics with respect to the multilevel impact of digital transformation as a way to provide recommendations for the strategic management of AI that will contribute to enhancing the long-term performance of organizations.

In section 2, we outline sociomateriality and the sociotechnical system theory as theoretical lenses. We then describe, in section 3, our methodology and its two stages (*conceptual stage* and *review stage*). Section 4 is an elaboration of our conceptual framework as a result of the first research stage, the *conceptual stage*. Section 5 presents the results from the second research stage, the *review stage*. This is followed by a discussion of our results in section 6. The paper concludes in section 7 with a summary, limitations, and an outlook.

Theoretical lens of sociomateriality and sociotechnical system theory

As spelled out in the Introduction, we ground our work in that of Leonardi (2012), who applies the discourse on sociomateriality in a conceptualization of sociotechnical systems. Sociomateriality melds technology and social phenomena to better describe how technology (as the material entity) is created and made observable through a variety of social phenomena (e.g., making decisions or formulating strategies as the social entity) that occur when the social and material entities are entangled in everyday practices. In contrast, sociotechnical systems describe the object of studying sociomateriality, such as organizations that are made up of the social (e.g., communication networks or status) and the material (e.g., technologies) (Leonardi, 2012). Sociomateriality draws on Actor-Network Theory (ANT), which can help further describe the power of the entanglement between the social and material entity. ANT considers how "network dynamics could … shape what people thought about a new technology, as well as whether and how they would use it" (Contractor et al., 2011, p. 683); using the lens of sociomateriality that builds upon ANT is appropriate for identifying interdependencies between humans (incorporating *virtue ethics*), technologies, and other non-human objects, such as management practices (incorporating *duty ethics*) based on how they are entangled and arranged in organizations and how those interdependencies lead to different forms of associations and power (Baron and Gomez, 2016; Contractor et al., 2011). It helps explain how those interdependencies affect actions, such as ethical or non-ethical behavior, that emerge in the collection of human and non-human objects (Cecez-Kecmanovic et al., 2014; Contractor et al., 2011; Latour, 2005). Hence, we use Leonardi‘s (2012) conceptualization to derive a theoretical lens grounded in sociomateriality, sociotechnical systems theory, and ANT.

Following Leonardi‘s (2012) conceptualization, an organization as a sociotechnical system comprises the *social subsystem* and *sociomaterial practices* (sometimes referred to as the *technical subsystem). Sociomaterial practices* describe the entanglement of humans and technology, conceptualized as the social entity and the material entity (Orlikowski, 2007; Orlikowski and Scott, 2008). While humans decide how to respond to technology based on their (human) agency (Cooren, 2004, p. 377; Leonardi, 2011), which also is influenced by their ethical ideology and being (*virtue ethics*), the material agency can be described as "the way objects act when humans provoke it" (Leonardi, 2012, p. 37); this also involves their ethical actions as *duty ethics*. Further, s*ociomaterial practices* shape and are shaped by "institutionalized ideas about how people could and should relate to one another" (Leonardi, 2012, p. 40), which can be described as the *social subsystem* (Leonardi, 2012). In the *social subsystem*, there can be several human objects (e.g., a manager's way of leading or the workforce's culture, incorporating *virtue ethics*) and non-human objects (e.g., guidelines, practices, strategies, or other technological objects, incorporating *duty ethics*) that affect the ethical management of human-AI interaction. Thus, with a collection of human and non-human objects as part of the *social subsystem* and the social entity and material entity as part of *sociomaterial practices*, the lens of sociomateriality enables us to improve our understanding of their "entanglement" (Cecez-Kecmanovic et al., 2014, pp. 825–826).

In the following, we stick to the terms of Leonardi (2012) by using "social entity" when considering the human interacting with AI and "material entity" when talking about AI interacting with the human in *sociomaterial practices*. In comparison, when talking about the *social subsystem*, we refer to human objects (e.g., a manager's way of leading or the workforce's culture) and non-human objects (e.g., guidelines, practices, strategies, or other technological objects).

To explain and outline differences in actions and behavior resulting from the interdependencies in an organization, we use the five notions of sociomateriality first introduced in the work of Orlikowski (2010, 2007) and Orlikowski and Scott (2008) and subsequently discussed by Jones (2014), which account for the complexity of human-AI interaction in the sociotechnical system theory:

i) *Inseparability* describes how the social and the material entities are interwoven;
ii) *Relationality* describes how the entanglement of the social and material entity affects both entities (social and material entity) as an outcome;
iii) *Performativity* describes how the social entity creates its reality;
iv) *Materiality* describes how the material entity is arranged based on an artifact's physical or non-physical objects; and
v) *Practices* describe how the entanglement is embedded in the *social subsystem.*

The notions of sociomateriality can contribute to our understanding of how ethical interdependencies arise between the *social subsystem* and *sociomaterial practices*. For instance, managers can act as role models, advising and guiding employees to make more ethical decisions when collaborating with AI acting as role models (Alsheibani et al., 2020; Eitel-Porter, 2021). The notion *performativity* enables us to explain how managers' *virtue ethics* (belonging to the *social subsystem*) are affected by how those managers are embedded in organizations, as well as how their ethical behavior (based on their *virtue ethics* on a collective level) might influence how employees as social entities change their own reality on (*virtue*) ethics in human-AI interaction, leading to ethical or non-ethical behavior. Hence, we adapt Leonardi's (2012, p. 43) conceptualization of sociotechnical systems by applying the five notions of sociomateriality (*materiality*, *inseparability*, *relationality*, *performativity*, and *practices*) to analyze the literature from our review.

Given that the interdependencies between the arrangement of all entities (material and social) collected in the sociotechnical system affect their actions and behavior (Contractor et al., 2011; Latour, 2005), we must further be aware of the strength of those interdependencies – mainly as AI as the material entity increasingly acts autonomously (Baird and Maruping, 2021; Rieder et al., 2020), changing the nature of human-AI interaction in terms of *sociomaterial practices* (Cecez-Kecmanovic et al., 2014; Maedche et al., 2019). To account for this changing nature of human-AI interaction, we follow the arguments of Jones (2014) and differentiate between *weak* and *strong* notions of sociomateriality.

In general, *weak* sociomateriality assumes a mutual interdependency between the social entity (humans) and the material entity (AI), meaning that each entity, with preexisting and independent attributes, forms, and capabilities, has its own properties that do not change (Cecez-Kecmanovic et al., 2014). This view is also represented in the substantialist ontology as one possible ontology of sociomateriality. It is supported by Faulkner and Runde (2011), Leonardi (2012), and Mutch (2013), who argue that the social and material entities exist as "separate and self-contained entities that interact and affect each other" (Cecez-Kecmanovic et al., 2014, p. 809). Using our aforementioned example of managers as human objects guiding ethical human-AI interaction in organizations from the *social subsystem*, a *weak* form of the notion *performativity* is present when a manager's behavior may lead to changes in the ethical behavior of an employee but may not affect their ethical ideology (*virtue ethics*). Hence, the individual's *virtue ethics* has independent attributes that neither change through the external influence of managers from the *social subsystem,* nor through the interaction with the AI as the material entity.

In comparison, *strong* sociomateriality assumes that the social entity (human) and the material entity (AI) are formed only through enactment. This is in line with relational ontology, another philosophy of sociomateriality that advocates that the social and material entities exist only in the context of that relationship, given that they are "inherently inseparable" (Cecez-Kecmanovic et al., 2014, p. 809). This view is supported by – among others –Barad (2003), Latour (1992), and Orlikowski (2007; Orlikowski and Scott, 2008), and draws on ANT, in particular, to focus on how the power of interactions depends on the association drawn by humans (Baron and Gomez, 2016; Contractor et al., 2011; Law, 1992). A *strong* form of the notion *performativity* is present when managers act, for instance, as role models, influencing not only employee ethical behavior but also their ethical ideology (*virtue ethics*) regarding interacting with AI. Further, changes in an individual's ethical ideology can depend upon various factors, such as influences through the interaction with AI as material entity or other influences originating from the social subsystem. Hence, using the *weak* and *strong* notions of sociomateriality allows us to better explain the strength of the interdependencies between the entities (social and material entity) that are arranged in the sociotechnical system.

To sum up, by differentiating between *weak* and *strong* notions of sociomateriality, we aim to provide a holistic and systematic view of the dynamics between AI (material entity), employees (social entity), and other human and non-human objects in the *social subsystem* – all incorporating the *virtue ethics* and *duty ethics* on an individual or collective level to some extent – that are arranged in an organization as sociotechnical system. This helps us to answer our research question regarding how we can understand and further explain ethical management of human-AI interaction from both the perspective of AI interacting with humans (SQ1), and that of humans interacting with AI (SQ2), as it allows us to outline the interdependencies of *virtue ethics* and *duty ethics* in the sociotechnical system. (Appendix A provides an overview of *weak* and *strong* notions of sociomateriality based on Jones (2014) and adapted toward our context of human-AI interaction.).

Methodology

Following the classification of literature reviews of Paré et al. (2015), we aimed to conduct a *theoretical review* grounded in a broad literature search and qualitative content analysis of relevant articles to elaborate and expand knowledge on the ethical management of human-AI interaction through the lens described above. Our research design comprises two stages: first, a *conceptual stage* (described in section 3.1) in which we developed deductively a conceptual framework to structure our review and aid in theorizing; and second, a *review stage* (described in section 3.2) encompassing the qualitative analysis of the literature identified based on the search protocol. Our analysis goes beyond synthesizing the literature by extending the main constructs from the conceptual stage and further examining their relationships through the notions of sociomateriality, which is the theorizing part of our review. In the following paragraphs, we elaborate on the two stages.

Conceptual stage

The aim of this stage is to derive a conceptual framework *deductively* to guide the theorizing part of our *theoretical review* in the second stage. We establish a deductive foundation that can be extended through the theorizing part of the review, adopting an *integrating* type of conceptualizing by following the guidelines for conceptual work proposed by MacInnis (2011). Integrating involves the synthesis of main constructs – ethics, human-AI interaction, and strategic management in our work – by taking what is known so far and transforming it into a novel structure (here integrating main constructs in sociotechnical system theory as conceptualized by Leonardi (2012)) by looking at the connections between all relevant constructs (here through the lens of sociomateriality) to gain a holistic perspective (MacInnis, 2011). It is important to note that as frameworks aim "to integrate" main constructs to define boundaries (Schwarz et al., 2007, p. 44), we do not claim comprehensiveness when synthesizing existing knowledge from relevant work of reference disciplines (Burton-Jones et al., 2021) for this integration. The integrative framework in the *conceptual stage* served as an "architectural plan" (MacInnis, 2011, p. 138) to guide the *review stage*. To provide an insightful framework with desirable quality, we followed the criteria of assessment of Schwarz et al. (2007). Appendix B details how we corresponded to these criteria. Our steps to develop the framework as an *integrating type* of conceptualization are described below.

First, we used our main constructs "ethics," "human-AI interaction," and "strategic management" as keywords to search for relevant literature in the databases listed in Appendix C. To identify existing work, we also used synonyms and similar concepts, such as "human-AI hybrids" (e.g., Rai et al., 2019) for the main construct of "human-AI interaction" and "AI governance" (e.g., Shneiderman, 2020) for the main construct "ethics." In our search, we used database filter-options to show the "most cited" articles (if available) and/or to sort articles with respect to their relevance. We also searched for current literature reviews summarizing the academic discourse on our main constructs. Our aim was to identify the most relevant articles in the academic discourse.

Based on the most relevant articles that addressed our main constructs, we conducted a backward search to check the references those articles used for their theoretical foundations (Weber, 2012). In particular, we checked whether there were similar or often-used references, following a snowball sampling strategy.

All articles identified as being highly relevant to our topic under research were used to identify and synthesize subconstructs that we employed in the next step to explain our main constructs in more depth. We identified 13 articles relevant for the synthesis of the subconstructs "substitution," "augmentation," and "assemblage" as different types of human-AI interaction; 16 articles remained for the synthesis of subconstructs representing principles of "ethics" (in specific, 9 articles representing business ethics and 7 articles representing AI ethics, being synthesized together as categories of "human-AI ethics"); and 9 articles remained for the synthesis of the subconstructs "management levels" (n = 4) and "management process" (n = 5) of the main construct "strategic management." Appendices M−O provide an overview of the identified literature we used to derive our subconstructs.

We next focused on literature that contributed directly to our research process and goals. Following Burton-Jones et al. (2021), we limited our analysis to the subconstructs that are focal to the understanding of our research subject:, we focused on literature contributing to our focus on the "strategic management" of human-AI interaction; and we used the subset of articles resulting from the snowball sampling that were most relevant to our theory development objective, as it is important to "identify the set of manuscripts that contribute to a theory and distinguish them from those that cite the theory for other purposes and be able to safely ignore those" (Larsen et al., 2019, p. 892). This narrowed down the literature to those subconstructs that are most relevant to establishing an *integrating type* of conceptualization: 8 articles remained to explain the types of human-AI interaction as subconstructs of the main construct "human-AI interaction"; 9 articles remained to explain the subconstructs representing principles of human-AI ethics of the main construct "ethics"; and 8 articles remained to explain the subconstructs of management levels and processes belonging to the main construct "strategic management.".

We then integrated our main constructs and their subconstructs into Leonardi's (2012) conceptualization of *sociotechnical systems*. To avoid "poor construct conceptualization" that could lead to misspecifications during theory development (MacKenzie, 2003, p. 324), we provided explanations for how the subconstructs are interrelated and how we embedded them into Leonardi's (2012) sociotechnical system. This resulted in a conceptual framework (see section 4) that served as the foundation for our *theoretical review*, in which we outlined how the notions of sociomateriality help us systematically understand ethics in human-AI interaction.

Review stage

We began the review stage by gaining initial insights from a preliminary literature review conducted between April and November 2021 and based on the terms from the main constructs detailed in the *conceptual stage*: types of human-AI interaction, ethics, and strategic management. The research team discussed the results and undertook conceptual mapping exercises to scope the literature in relevant fields. This process served as a stage-gate to the review process and yielded the results required for developing the search strategy. We adapted our framework based on the results of this preliminary literature review, relying on iterative rounds and leading to our research design, thus contributing to the quality of the outcome of theorizing (Rivard, 2021).

Next, we followed the steps suggested by Templier and Paré (2018) for conducting a *theoretical review*. We assessed the seven databases that index scientific publications relevant to our review to ensure broad coverage of interdisciplinary literature: *ACM Library* (ACM), *AIS Electronic Library* (AIS), *Business Source Ultimatum via EBSCOhost* (EBSCO), *IEEE Computer Society Digital Library* (IEEE), *JSTOR*, *ScienceDirect* (SciDir), and *Web of Science* (WoS). The databases cover major academic journals from the domains of Information Systems (IS), Management and Organizational Science, and ethics (see Appendix C). We manually checked leading conferences and journals that are not included in these databases (see Appendix D).

We conducted a search based on the keywords derived from the first stage (Kitchenham and Charters, 2007). The primary keywords are derived from the constructs of our conceptual framework: the types of human-AI interaction and ethics in the context of strategic management. In addition, we adopted four abstract search terms that refer to the lines of research relevant to our work: "AI," "strategy or management," and "ethics" (see Appendix E). This increased the number of potentially relevant results, allowing us to gain a broader picture of the current literature (Boell and Cecez-Kecmanovic, 2015). We included articles irrespective of their publication date.

The search was carried out in June 2022 and updated in January 2023. It yielded 10,009 hits (9,136 hits in the main search and 873 hits in the update) across all databases without the use of database filter options; 6,904 hits (6,058 hits in the main search and 846 hits in the update) across all databases with the use of database filter options (see *format screening*); and 6,377 unique hits (5,688 hits in the main search and 689 hits in the update) after the elimination of 527 duplicates (n = 370 in the main search and n = 157 in the update). We added 1 hit through a manual search in the five leading conferences and journals not covered in the databases. In total, we received 6,905 overall hits (with database filter options); 6,378 unique hits remained for the subsequent manual screening processes.

In the next step, we applied exclusion and inclusion criteria through *format*, *quality*, and *relevance screenings*. In the *format screening*, we applied several exclusion criteria to remove non-scientific literature, incomplete articles, and gray literature (see Appendix F). We had already applied format criteria through database filter options whenever possible (see preceding step) and thereby removed 3,105 articles. Since this option is not available for all outlets, the format criteria also had to be checked manually (see Appendix F). In total, 2,379 articles were eliminated manually due to format criteria, and 3,999 articles remained.

We conducted a further *quality screening* based on various established rankings to restrict our review to high-quality publications that define the state of the art and to reduce the set of papers considered to a quantity appropriate for qualitative content analysis (see Appendix G). During the preliminary literature review, we identified several ethics journals that are important to the community but not present in any ranking. As an exception to the *quality screening* process, we included these journals to achieve a balance between interdisciplinary research streams and advance diversity in developing theory (Markus and Rowe, 2021). At the end of the *quality*

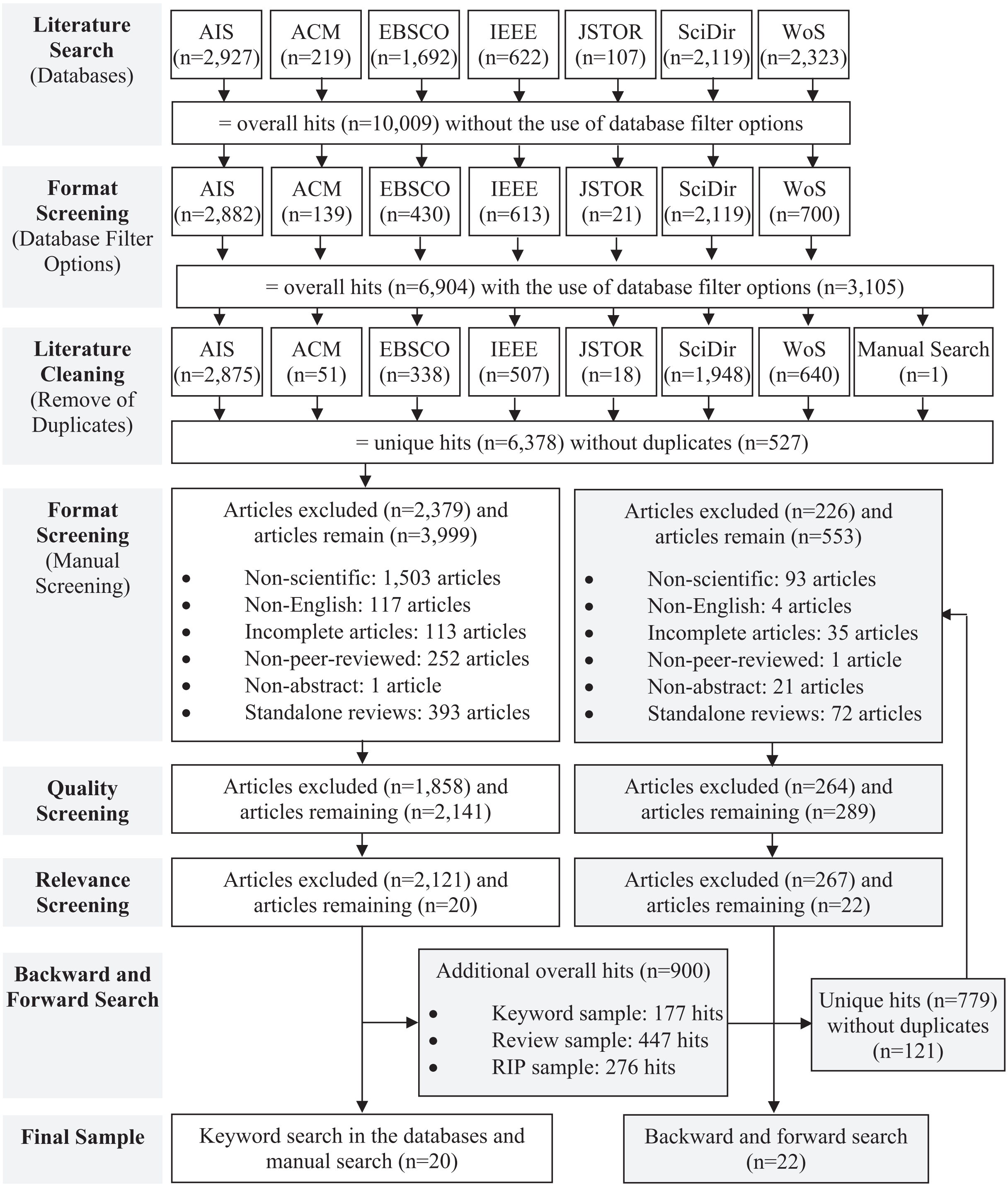

Fig. 1. Search Process, based on Templier and Paré (2018).

screening 1,858 articles were removed and 2,141 remained.

Finally, we performed several rounds of *relevance screening* to increase the quality of our results. In the first round, we chose a small subsample of the articles (n = 133) randomly for the parallel training purposes of the coders. The relevance criteria were revised through subsequent discussion. We used Cohen's Kappa, following Cohen (1960), to measure interrater agreement and increase our review reliability, obtaining a score of $\alpha = 0.81$, which indicates a strong agreement between coders (McHugh, 2012). Next, each of the coders coded an equal share of the remaining articles. In a third round, the articles were exchanged between the coders, thus ensuring that codes were double-checked to reach a strong level of reliability.

The *relevance screening* process was based on the full content of each paper. To be deemed relevant, each paper had to meet a set of three criteria. First, articles had to provide contributions regarding ethics in the context of human-AI interaction. Hence, for example, articles that analyze AI design principles rather than focusing on the human-AI interaction were excluded; while such articles may answer questions about how we might understand ethical AI development, they do not typically address ethical management of AI interacting with humans (see SQ1). Second, the AI-based technology mentioned in the articles had to carry out a task that typically requires human cognition, following the definition of AI based on Rai et al. (2019). Third, to be considered relevant, articles must have included ideas on the management of human-AI interaction in an organizational context (especially due to SQ2, which aims at understanding the management of humans interacting with AI). See Appendix H for a complete list of criteria. In the end, 2,121 articles were excluded, and 20 articles remained.

Since a keyword-based search may not identify all relevant articles, we also conducted a *backward and forward search*. For the *backward search*, we checked the sources cited in each of the 20 articles that remained in our final sample after the keyword search in databases and manual search (keyword sample). In addition, we conducted a *backward search* and *forward search* for the previously excluded literature reviews (review sample) and research-in-progress articles (RIP sample). For the *forward search,* we looked for articles that refer to articles of the three samples: keyword, review, and RIP. The articles identified through the *backward* and *forward searches* were further checked against the inclusion and exclusion criteria described above. The remaining articles were subjected to further rounds of *backward* and *forward searches*, a process repeated through several rounds until it no longer yielded relevant articles not already present in the dataset (see Appendix I). This added another 22 articles to our final sample.

Overall, the systematic literature search yielded a *final sample* of 42 relevant articles meeting the inclusion criteria for format,

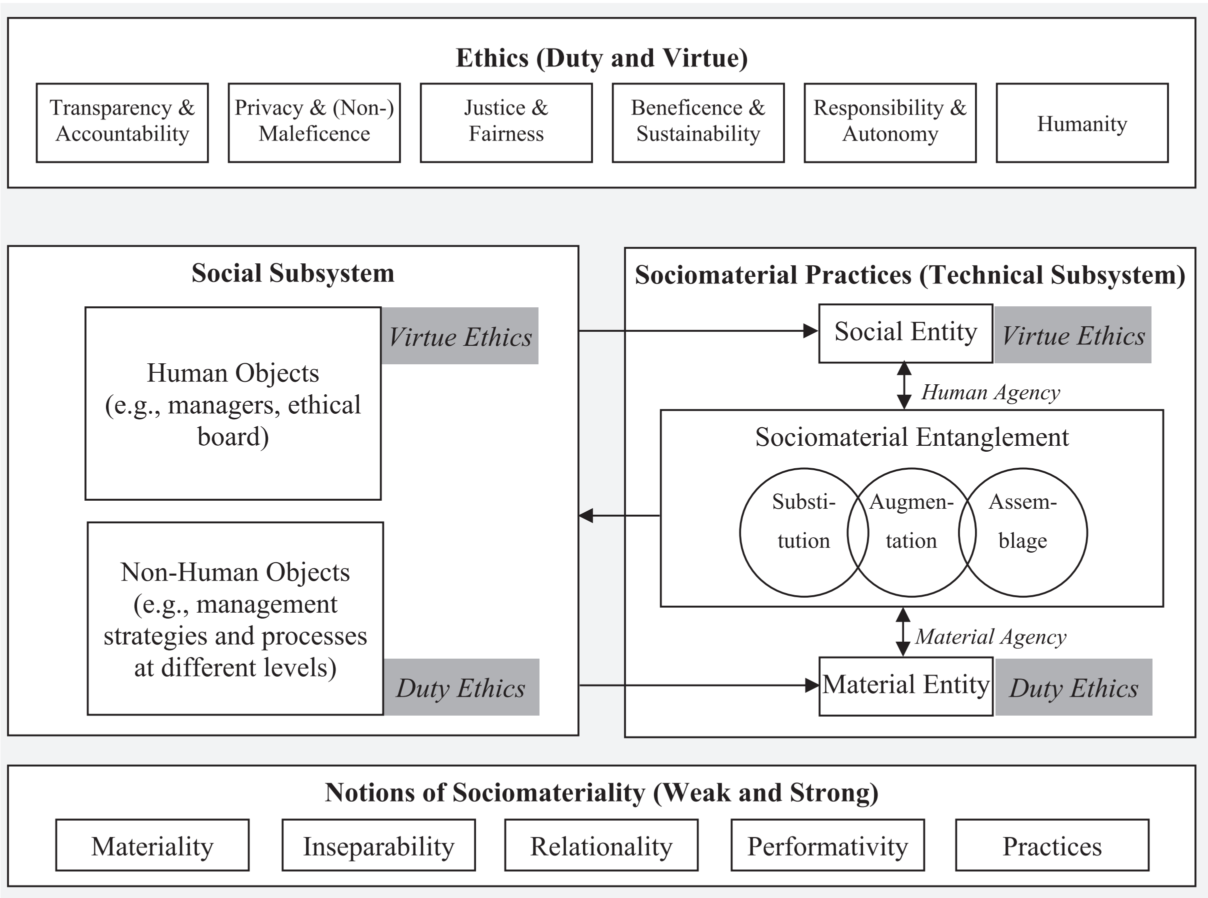

Fig. 2. Conceptual Framework of Ethical Management of Human-AI Interaction.

quality, and relevance across all databases and backward and forward searches (see Appendix J). Before downloading the PDFs of our final sample into the software MAXQDA for qualitative analysis, we structured the sample's information as *meta information* to provide a descriptive overview of our final sample (see Appendix K). Meta information describes the articles based on their main characteristics (e.g., research domain, method) that differentiate between different article types.

Fig. 1 is an overview of the search process based on the steps suggested by Templier and Paré (2018).

For the *qualitative analysis* of the final sample of articles, we structured the content of relevant articles according to the constructs comprising our deductively derived framework (Mayring, 2014, pp. 95–98). To guarantee consistency during the coding process, we used a codebook that details all deductively derived codes in terms of definitions, anchored examples, and coding rules for each category, as suggested by Mayring (2014, p. 95). The codebook was developed by all coders, and its reliability was improved through discussions among the coders and repeated coder training. Appendix L shows the coding scheme. The deductive analysis reflects the *extend strategy* of "Next-Generation Theorizing" of Burton-Jones et al. (2021) by increasing our understanding of the transforming nature of human-AI interaction as the changing phenomenon through the use of the notions of sociomateriality in the context of sociotechnical system theory (as well-established theories), and by extending those theories.

Conceptual framework

The following pages explain the three main constructs (ethics, human-AI interaction, and strategic management) and their sub-constructs that are relevant to increasing our understanding of ethical management of human-AI interaction. We further integrated them into Leonardi's (2012) conceptualization of sociotechnical systems to demonstrate initial interdependencies. This process resulted in a deductively derived conceptual framework (see Fig. 2) that serves as a foundation for the theorizing part of our *theoretical review* and helps us to better reflect upon the strategic management of human-AI interaction from both perspectives: managing AI interacting with humans (SQ1), and managing humans interacting with AI (SQ2).

As Fig. 2 illustrates, we identified three main constructs as relevant to our research context:

i) The types of human-AI interaction (*substitution*, *augmentation*, and *assemblage*) are part of *sociomaterial entanglement* and are described further in section 4.1.
ii) Ethics (especially *virtue ethics* and *duty ethics*) can be part of either the *social subsystem*, representing ethics on a collective level, or *sociomaterial practices*, representing ethics on an individual level. This is detailed in section 4.2.
iii) Strategic management represents the context in which human-AI interaction can be managed and is part of the *social subsystem* (see section 4.3).

The arrows in Fig. 2 between the *social subsystem* and *sociomaterial practices* are based on Leonardi (2012), who promotes the substantialist ontology (*weak* sociomateriality) and suggests that the *social subsystem* and *sociomaterial practices* are mutually dependent. *Sociomaterial entanglement* may affect the *social subsystem*, while processes within the *social subsystem* "can shape future patterns of imbrication, [that] can in turn bring changes to an artifact's materiality or a person's intentionality" (Leonardi, 2012, p. 43). Even though we adopt this ontological view in the conceptual framework of Fig. 2, we follow the recommendation of Jones (2014) in our second stage, when theorizing based on the literature review, to achieve a more differentiated analysis of *weak* and *strong sociomateriality* through the notions of sociomateriality.

How the types of Human-AI interaction refer to sociomaterial entanglement

As Fig. 2 outlines, human-AI interaction resembles *sociomaterial entanglement* embedded in *sociomaterial practices*. Since increased capabilities of AI lead to changing natures of sociomaterial entanglement, we need to differentiate between different types of human-AI interaction. It can be classified into three archetypes (see Appendix M): *substitution*, *augmentation*, and *assemblage*.

Substitution refers to AI that replaces humans in performing a task (Rai et al., 2019). Strictly speaking, no specific human-AI interaction concerns the specific task the AI conducts. Hence, ethical management must focus on the conditions surrounding these tasks, as well as upstream and downstream processes.

Augmentation focuses on improving each other's abilities (Rai et al., 2019) and balancing each other's weaknesses (Daily et al., 2017; Raisamo et al., 2019). Either the AI can improve the human's performance (*human augmentation*; Wilson and Daugherty, 2018, p. 214), or the human can improve the AI's performance (*machine augmentation*; Mnih et al., 2015). We feel, though, that this definition does not adequately reflect different cases of augmentation, as there may be different outcomes (positive or negative). In some cases, humans as the social entity may try to exploit the algorithm as the material entity to increase their personal benefits (Lee, 2018; Lee et al., 2015; Zhang et al., 2020), as awareness of the weaknesses of the systems surrounding the AI makes humans able to behave opportunistically (representing a rather egoistic form of opportunistic behavior). In other cases, humans may "outsmart" the AI, not for their own benefit, but rather for an overarching goal from which others benefit (representing rather an altruistic form of opportunistic behavior). In this way, one entity may try to balance the insufficient moral behavior of another in accordance with ethical norms. The ideology of the social entity does not subordinate morally correct behavior to one's own advantage; instead, it is about weighing different moral maxims against each other (e.g., helping others vs. behaving according to transparency standards). Thus, when considering managing *augmentation*, we need to account not only for the behavior but also the ethical intention of humans (that may be positive or negative) that guide the behavior in human-AI interaction.

Assemblage (often referred to as hybrid intelligence) suggests that combining humans and AI agents in an integrated unit to perform complex tasks leads to better outcomes through collective intelligence (Rai et al., 2019). We also suggest patterns of interpreting *assemblage* as a process of several augmentative or substitutive tasks being executed sequentially or in parallel, leading to “superior results” (Dellermann et al., 2019a, p. 19). This process view demonstrates that there can be tasks during human-AI interaction in which the AI replaces the human and other tasks in which the human or the AI is augmented by the other entity.

Fuegener et al. (2021, p. 32) state that “metaknowledge” is required to achieve “superior results” with AI by deciding which tasks should be delegated to the AI (*substitution*) and which should remain with humans (*augmentation*), given that “poor delegation decisions” can lead to “unconscious traits” and a low level of performance. Bansal et al. (2021, p. 13) state that “in an ideal team, the human and AI would have minimally overlapping mistakes so that there is a greater opportunity to correct each other’s mistakes.” This demonstrates that *assemblage* is built upon *substitution* and *augmentation*. Therefore, ethical management of assemblage needs to consider both upstream and downstream processes (resulting from the insights we derived for substitution) as well as ethical behavior and the intention of the social entity (which may be positive or negative), which we learned from the insights we derived for augmentation.

While we adopted different types of human-AI interaction as defined by Rai et al. (2019), we aim to create awareness that ethical management of human-AI interaction cannot be restricted to a specific task and, thus, to a specific type of interaction. Consequently, the process of entanglement in the context of *sociomaterial practices* cannot be strictly classified according to these types of human-AI interaction. Instead, these archetypes should be understood as tools that help us make the process view and its effects on ethical management more tangible. We refer to this in the second stage when presenting and discussing the review results.

How ethics is considered as part of both the social subsystem and sociomaterial practices

As Fig. 2 illustrates, ethics on an individual level can be subordinated to the social entity and the material entity being part of *sociomaterial practices*, but also on a collective level to human objects and non-human objects being part of the *social subsystem*. Below, we deepen our understanding of ethics and its subconstructs to understand the differences in ethical considerations that can be drawn depending on where in sociotechnical systems we are talking about ethics.

In general, ethics can be seen as “the capacity to think critically about moral values and direct our actions in terms of such values.” (Churchill, 1999, p. 259). Hence, ethics is thinking not only about moral values as “the intention of the person making the decision or performing the act,” but also about “the consequences of … an action or decision” (Hosmer, 1987, pp. 321–322). In this work, we refer to *virtue ethics* when referring to a person’s ethical intention, which is guided by a person’s ethical ideology and characterized by the person’s moral values and assumptions determining the intrinsic motivation to behave ethically or not (Eitel-Porter, 2021; Flathmann et al., 2021; Gandz and Hayes, 1988). Beyond *virtue ethics* that guide ethical behavior, there are also principles, rules, norms, and standards that guide ethical behavior by obligation, which we refer to as *duty ethics* – also known as deontology (Bowie, 1998; Flathmann et al., 2021; Jones et al., 2005; Kant, 1991, pp. 66–70).

As outlined in the Introduction, ethical management of human-AI interaction on a strategic level requires consideration of both ethical perspectives, *duty ethics* and *virtue ethics*, both of which are represented in the *social subsystem* and *sociomaterial practices*. *Duty ethics*, inter alia, is part of the social subsystem, as it reflects guidelines and norms in an organization (Leonardi, 2012, p. 43). The *social subsystem* also includes the *virtue ethics* perspective since the workforce has its own ethical understanding based on collective moral values that drive ethical behavior in the organization regardless of *duty ethics*, and that might differ from the ethical ideology of any individual worker (Coldwell et al., 2008; Stevens, 1999). Both ethical perspectives are also part of *sociomaterial practices*, given that within their enactment, the AI as the material entity is designed on ethical principles guiding the material agency (*duty ethics*), and the employee as the social entity has its own inherent ethical ideology guiding human agency (*virtue ethics*) (Flathmann et al., 2021; Siau and Wang, 2020). While the *social subsystem* comprises ethical principles and guidelines on a collective level (*duty ethics*) and corporate values based on the ethical understanding of the workforce (*virtue ethics*), *sociomaterial practices* comprise both ethical perspectives on an individual level concerning a specific task.

When applying both ethical perspectives to our underlying research subject, we also have to consider established research in related domains, specifically AI ethics and business ethics. AI ethics corresponds to SQ1 (AI interacting with humans), whereas business ethics is linked to SQ2 (humans interacting with AI). Both AI ethics and business ethics are based on principles that guide ethical behavior (Floridi et al., 2018; Rezaee, 2008, p. 64) (see Appendix N). AI ethics comprises “a set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies” (Leslie, 2019, p. 3). Business ethics comprise organizational norms and guidelines that are required to guide ethical behavior among organizational stakeholders, as some people in organizations may act in what could be characterized as a morally insufficient manner (Jones et al., 2005; Parker, 2002). AI ethics and business ethics correspond primarily to the *duty ethics* perspective that aims to foster ethical behavior in human-AI interaction and, to some extent, aims to strengthen an individual’s ethical understanding (what we refer to as *virtue ethics*).

To make ethical considerations in human-AI interaction more tangible, we draw upon the principles discussed in the literature on AI ethics and business ethics and derived six categories of “human-AI ethics” (for details, see Appendix P):

i) *transparency & accountability* aim to make human-AI interaction processes easy to understand and decisions reviewable, and to ensure that processes of human-AI interaction can be investigated depending on the interests of both entities (social and material entity);
ii) *privacy & (non–)maleficence* aim to protect personal privacy and security in human-AI interaction, avoiding misuse of AI through humans and vice versa;

iii) *justice & fairness* aim at respecting all the interests of both entities (social and material entity) to maximize inclusion and fair treatment in all types of human-AI interaction;
iv) *beneficence & sustainability* aim at beneficial human-AI interactions with a long-term positive impact;
v) *responsibility & autonomy* aim to balance the power of AI and humans in their interactions and manage human-AI interaction based on the agreement among all entities (social and material entity) involved or affected by its consequences; and
vi) *humanity* aims to maintain social values and foster social togetherness in the context of human-AI interaction

We refer to these categories of human-AI ethics to better explain the interdependencies between the *virtue* and *duty* ethics perspectives in managing human-AI interaction.

How strategic management represents the context of the social subsystem

In the context of information technologies (IT), a strategy concerns "the ability to identify and evaluate the implications of IT-based opportunities as an integral part of business strategy formulation" (Peppard and Ward, 2004, p. 176). Strategic management includes strategies at three levels. At the *corporate level*, strategy (corporate or global strategy) concerns the strategic planning and coordination of several substrategies (business strategies). At the *business level*, a strategy aims to gain a competitive advantage and achieve the organization's overarching goals and vision (Karimi and Konsynski, 2002, p. 92; Ritson, 2008, p. 13; Somogyi and Galliers, 2009, pp. 11–12). The *functional level* has strategies based on different departments, such as accounting and HR (Karimi and Konsynski, 2002, p. 93; Somogyi and Galliers, 2009, p. 15). To achieve a "strategic fit" (Henderson and Venkatraman, 1999, p. 476), it is important to align these strategies with the organization's processes.

As Fig. 2 illustrates, strategic management can be interpreted as a context being part of the *social subsystem* that guides ethical behavior through strategies and processes before (preventive), during, and after (reactive) the human-AI interaction in an organization. Strategic management establishes ethical principles that provide guidance and standards for how human-AI interaction can be accompanied by efforts to manage it ethically. When managed well, AI can increase the quality of work, but poorly managed AI can increase opacity and inequality in work arrangements (Faraj et al., 2018; Jobin et al., 2019). The strategic management of human-AI interaction in an organization can be split into four processes: *planning, strategic*, *tactical*, and *operational processes* (see Appendix O for an overview). Given the relationship between strategic management and organizational goals (Henderson and Venkatraman, 1999; Peppard and Ward, 2004), organizations must involve their ethical understanding when formulating strategies and aligning their processes to achieve ethical management of human-AI interaction. This means that an organization's ethical principles are represented in certain ways through its strategies and processes – what can be referred to as AI governance (Gasser and Almeida, 2017) and corporate governance (Bonn and Fisher, 2005; Rezaee, 2008). For example, to enforce principle-based ethics, organizations may follow external guidance (Rezaee, 2008; Shneiderman, 2020).

In the analysis that follows – the second stage of our theory-developing review – we aim to provide deeper insights into the ethical

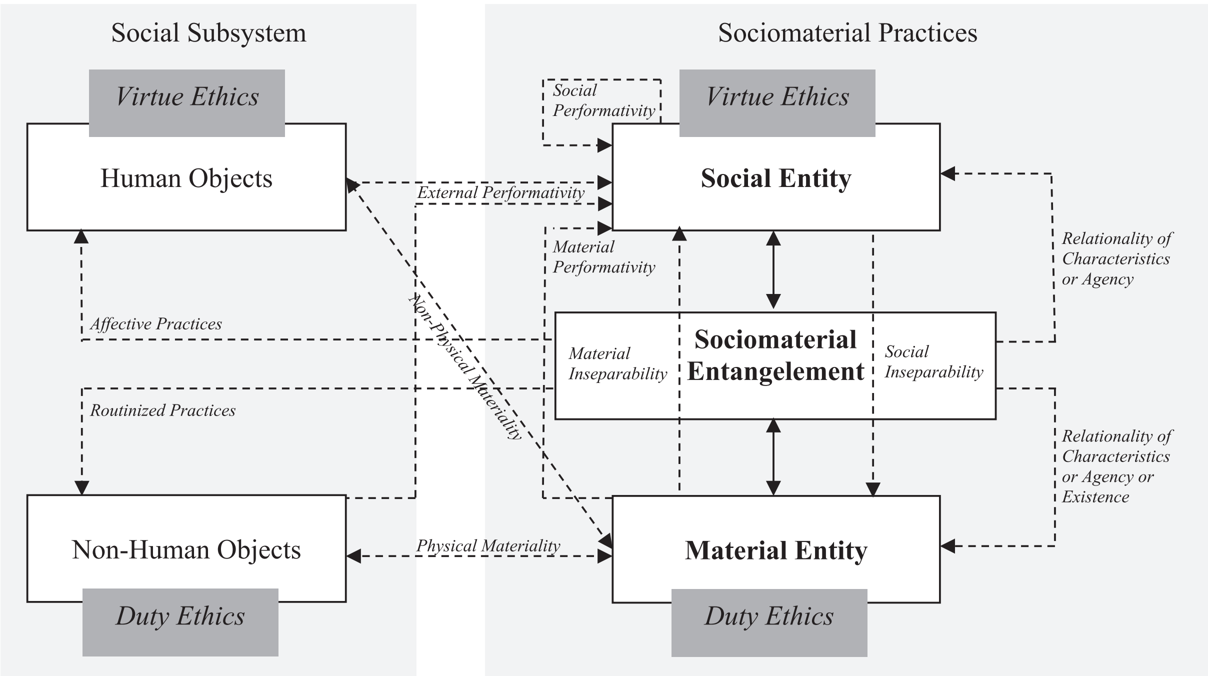

Fig. 3. Notions of Sociomateriality Determining Ethical Management in Sociotechnical Systems.

management of human-AI interaction by explaining the relationships and interdependencies between all constructs and their subconstructs we introduced in this stage through the five notions of sociomateriality.

Analysis and results

In this section, we describe and explain the main findings of our review along the five notions of sociomateriality (*inseparability*, *relationality*, *performativity*, *materiality*, and *practices*). The notions enable us to better reflect the complexity of human-AI interaction (*sociomaterial entanglement*) and explain the interrelationships with the corresponding management practices and processes (*social subsystem*) occurring in organizations as sociotechnical systems. Given that our underlying aim is to increase understanding of the ethical management of human-AI interaction through the lens of sociomateriality, we explain in this section how *virtue ethics* and *duty ethics* can constitute the ethical management of human-AI interaction corresponding to our two subquestions concerning AI interacting with humans (SQ1) and humans interacting with AI (SQ2). Hence, we concur with Jones‘ (2014) arguments regarding the relevance of using the potential of the notions of sociomateriality to explore sociotechnical interactions – especially when it comes to ethical considerations.

Fig. 3 is an overview of how the five *weak* and *strong* notions of sociomateriality can be integrated into the conceptual framework we derived in the conceptual stage (see section 4) based on the main findings of our literature review.

Although Fig. 3 represents the two ethical perspectives (*duty ethics* and *virtue ethics*) assigned to different entities (social or material entities) or objects in the sociotechnical system, there are no clear dividing lines between those two perspectives. For instance, Coombs et al. (2021, p. 4) provide an example that demonstrates these mutual interdependencies of (*virtue* and *duty*) ethics, given that "[f]or machines to be valuable overall, they must support higher-level, more socially beneficial thinking and discussion by humans" – and hence it is important to integrate *virtue ethics* in the AI's logic. Thus, the machine's behavior is based on "the lens through which [AI designers] look at and reason about the world is shaped by [their] conscious and, more importantly, unconscious values and assumptions" (Ehsan and Riedl, 2020, p. 12). Following this argument, the AI also inherits the developer's *virtue ethics* to some extent, resulting in programmed rules (which we refer to as the AI's *duty ethics*).

In the following, we detail the interweaving of all constructs constituting the ethical management of human-AI interaction and their interdependencies by investigating each notion of sociomateriality and introducing the subforms outlined in Fig. 3 based on the results of our literature review.

Inseparability: How interdependencies between the social entity and the material entity affect ethical management of human-AI interaction.

The notion of *inseparability* focuses on the interaction (moment of "imbrication") of the social and material entity (*sociomaterial entanglement*). We suggest distinguishing between two subforms of *inseparability*:

i) *Social inseparability:* the material entity depends on the social entity within the sociomaterial entanglement (e.g., decisions of humans build the material entity and how it operates). In a *strong* form, this can be aligned with *machine augmentation*, whereas in a *weak* form, it can be linked to *substitution*.
ii) *Material inseparability:* the social entity depends on the material entity within the sociomaterial entanglement (e.g., the material entity determines working practices). In a *strong* form, this can be aligned with *human augmentation*, whereas in a *weak* form, it can be linked to *substitution*.

The following two sections provide examples from our literature review to bolster our suggestion to differentiate between the two subforms and demonstrate how they shed light on the interdependencies between the social and material entity when directly entangled (wherefore we also relate the subforms to the types of human-AI interaction).

Weak social inseparability (the material entity operates without or with less human intervention) is linked to ***strong material inseparability*** (the work practices of the social entity are highly dependent on the material entity): In the context of recruitment, the AI (material entity) that can, for instance, reject job applicants without human intervention (Acikgoz et al., 2020) indicates *weak social inseparability*, as the AI decides independently of the recruiter (social entity); this indicates *substitution*. However, in the "second stage of staffing," the recruiter (social entity) depends on the AI's decision regarding whom to invite for a job interview, demonstrating that certain tasks cannot be completed without the material entity (*human augmentation*) and, thus, indicating *strong material inseparability*.

Strong social inseparability (the material entity cannot interact or operate without the social entity) is linked to ***weak material inseparability*** (the material entity falls out of the awareness of the social entity in work practices): In the context of loan decisions, the loan consultant – as the social entity having expertise on loan decisions – might not depend as much, compared to a non-experienced employee, on the AI-based systems (material entity) providing decision criteria for loan acceptance or rejections. Strich et al. (2021, p. 316) offer an example from an interview with a former loan consultant: "I know how to deal with the customer if the system is offline. It may sound mean, but the new colleagues would have a big problem, because they are only used to working according to the predefined path [given by the AI-based system]. In addition, if something beyond that happens, they would be helpless." In the example, the former loan consultant (social entity) does not depend on the AI (material entity), which indicates *weak material inseparability*. At the same time, the AI-based system depends on previous inputs of the social entity: the "AI system uses past data from similar customers profiles to predict future customer needs" (Strich et al., 2021, p. 315). That engenders *strong social inseparability* and refers to *machine augmentation*.

However, the level of interdependence can vary over time given that – following the aforementioned example of integrating an AI system for loan decisions – new colleagues are not experienced loan consultants (social entity). They may not be able to work without the AI (material entity) due to their lack of knowledge (Strich et al., 2021), which demonstrates a *strong material inseparability* and *human augmentation*. These interdependencies over time confirm the relevance of a process view when it comes to different types of human-AI interaction (see section 4), and so it is important to differentiate between several forms of *inseparability*. In the following, we elaborate on these four different forms of the notion *inseparability* and highlight how they affect the two different ethical perspectives (*virtue* and *duty*) in the management of human-AI interaction.

Beginning with ***weak social inseparability***, we recognize that this form is characterized by *transparency & accountability* issues (referring to the *duty ethics* perspective), given that "employees [note] a lack of control over the interaction and outcome" (Rix and Hess, 2022, p. 9) due to the complexity of the algorithms (material entity). This leads to "highly unpredictable" results (Mayer et al., 2020, p. 248) and "heightens [individual] perceptions related to the creepy nature of the [AI-based] technology" (Rajaobelina et al., 2021, p. 2351). Hence, a "clear capability to explain intent and process" is required (Someh et al., 2022, p. 151) to enhance the decision confidence of humans having to work with the AI's outcome (Liao et al., 2020). Hence, *duty ethics* is especially relevant when the AI (material entity) operates without human interventions (thus, no ability to involve someone's *virtue ethics*).

By contrast, ***weak material inseparability*** is linked to the *virtue ethics* perspective of humans since "usually any final decision … is still taken by humans – and in many cases, AI suggestions … are not implemented" (Klumpp and Zijm, 2019, p. 269). For instance, in the context of using an AI agent to predict a project's success or failure, "humans are using the agent as a filter [for] a first check" (Basu et al., 2021, p. 8). Not following the AI suggestions can be a "sign of resistance" or result from the human assuming the possession of "better knowledge" compared to the AI (Klumpp and Zijm, 2019, p. 269). Yet, the interaction can also promote ethical behavior when the material entity and the social entity operate as separate units (*weak material inseparability* in *substitution*). This is especially the case when we look at *beneficence & sustainability* from a *virtue ethics* perspective, given that AI can help create favorable work environments and increase work-life quality (Fuegener et al., 2021; Rix and Hess, 2022). For instance, the AI (material entity) can take over "disliked and routine tasks," leaving employees (social entity) to "focus on more value-creating and experimental tasks that correspond with their personal interests" (Rix and Hess, 2022, p. 7) and that are more related to "creative thinking" (Malik et al., 2021, p. 344). This represents a potential to enforce beneficence between humans and AI, interacting mainly independently of each other. However, as the interaction between humans as social entities is not mediated by the material entity, there may be other mechanisms required to ensure correct ethical behavior of both entities.

Concerning ***strong social inseparability***, we recognized the increasing importance of the *virtue ethics* perspective, as the material entity cannot operate adequately without the social entity. This can be linked mainly to the issue of *humanity*. Therefore, it is essential to align human values (*virtue ethics* of the social entity) with algorithmic values (*duty ethics* of the material entity) (Mayer et al., 2021). In this way, ethical guidelines (*duty ethics* of the *social subsystem*) can help create a "collective understanding of AI ethics in the corporation" (*virtue ethics* of the *social subsystem*) (Mayer et al., 2021, p. 6). The ethical guidelines of the *social subsystem* can moderate the alignment between humans' *virtue ethics* and the AI's *duty ethics*.

It is important, however, to include human judgment and the ability to create situational awareness in human-AI interaction (e.g., Malik et al., 2021, Park et al., 2021, Someh et al., 2022). For instance, one interviewee of Someh et al. (2022, p. 153) argues that "[i]f something is going against a person, we wanted humans to intervene and make final judgment as opposed to a machine making the judgment." Lee et al. (2015, p. 1610) demonstrated that there can arise "many complications … when one relies too heavily on quantified metrics without deeper considerations of their meaning and nuances." Hence, to foster humanity and reduce the emotional burden during human-AI interaction, it is important to align AI's *duty ethics* with human *virtue ethics*.

In comparison, ***strong material inseparability*** is often linked to ethical issues from the *virtue ethics* perspective regarding *responsibility & autonomy*. For instance, the "loss of critical thinking and knowledge" can become an issue for organizations in the long term, given that "employee's awareness is severely impaired or even nonexistent" (Mayer et al., 2020, p. 249). Thus, they simply transfer their responsibility to the AI, which leads to employees not justifying decisions (Strich et al., 2021). The loss of autonomy of the social entity further spurs *strong material inseparability* and can even increase when the AI restricts humans' critical thinking about decisions through a programmed carrots-and-sticks mechanism (*duty ethics*). For instance, the platform Uber, which automates "the exercise of managerial control, uses "material benefits and privileges through Uber Pro, as well as symbolic badges" following strict instructions of the algorithm without questioning it (Pregenzer et al., 2021, p. 9). It may, as a consequence, might punish people who criticize the decisions, which may lead to an increased loss of autonomy, leading to unintended negative consequences. Therefore, it is "essential that an individual who may be subject to a purely automated decision has the right and power to challenge the decision" (Polyviou and Zamani, 2022, p. 9). This is why it is important to establish not only adequate principles guiding human-AI interaction but also strengthen ethical understanding and behavior by enabling individuals to criticize AI's behavior.

To sum up, the direction of interdependencies (whether we have *weak* or *strong material inseparability* or *social inseparability*) may also influence ethical potentials or ethical issues that arise in human-AI interaction.

Relationality: How sociomaterial entanglement affects the existence of the social entity and the material entity or their characteristics, and thus, the ethical management of human-AI interaction.

Similar to the notion *inseparability*, *relationality* is also related to sociomaterial entanglement, but it focuses on the outcome of the human-AI interaction process as it describes how social and material entities, their existence, and their agency change through the entanglement. Since *relationality* focuses on the outcomes of human-AI interaction, it can be observed in all types of human-AI interaction. Yet, the same type of human-AI interaction can lead to different outcomes regarding the characteristics of the social or material entity, its agency, and the existence of other objects. We suggest distinguishing between three subforms of *relationality*:

i) *Relationality of characteristics*: certain characteristics of the entities (social or material entity) arise through sociomaterial entanglement (e.g., capabilities, attributes, forms).

ii) *Relationality of agency*: sociomaterial entanglement changes the agency of the entities (material or social; agency is the capacity to carry out an action following a cognitive process linked to the entity's intentions).
iii) *Relationality of existence*: the existence/presence of an entity is the product of sociomaterial entanglement.

The following two sections provide examples to bolster our suggestion to differentiate between the different subforms. After we introduce the difference between relationality of the social and the material entity, we explain the three subtypes by outlining how the outcome of *sociomaterial entanglement* can affect social or material entities.

Beginning with ***relationality of the social entity***, the implementation of an AI-based application system may remove "consultants from decision making" and thus affects "their professional image and external perceptions of their role" (Mayer et al., 2020, p. 248), which can be linked to *strong relationality of agency* (meaning that the agency of the social entity changes through the interaction). In the long term, the consultant's decrease in decision-making power can further lead to a change or decrease in expertise (Mayer et al., 2020), which can be linked to *strong relationality of characteristics* (the characteristics of the social entity change through the interaction). With respect to *relationality of existence*, new roles for the social entity can arise, given that "[h]uman-AI partnering will also require workers to manage shifts between micro-identities (i.e., new roles that are enacted in relation to the needs and requests of [AI-based digital assistants], and traditional human-facing roles)" (Cranefield and Doyle, 2022, p. 13). This is supported by Grønsund and Aanestad (2020, p. 1), who stated that "[i]n contrast to replacing human work, the emergent configurations required new roles and redistribution of extant expertise to augment and improve the accuracy of the algorithm." Due to these interdependencies, relationality of existence manifests primarily in *strong* relationality.

When it comes to ***relationality of the material entity***, we can see that through "human–machine collaboration, … machines become increasingly intelligent" (Jia et al., 2022, p. 388), as for example, feedback loops during the interaction "improve AI performance" (Liao et al., 2020, p. 6). This increased intelligence can be associated with a change in the *characteristics* of the material entity. The relationality can be either weak or strong, with the characteristics changing either through the human-AI interaction (indicating *strong relationality*) or existing independent of the human-AI interaction (indicating *weak relationality*). Concerning *relationality of existence*, Coombs et al. (2021, p. 5) state that the human-AI interaction may also result in "[n]ew roles for machines through AI," which indicates that the new roles' *existence* depends on the interaction representing a *strong* form of *relationality of existence*. Another example illustrates *relationality of agency*. When failure is an outcome of human-AI interaction, individuals sometimes "attribute [their] agency to the technology when they lost the game, compared to the success scenario" (Jia et al., 2022, p. 393). In doing so, the social entity transfers the agency of being responsible for the outcome to the AI in either a *strong* or a *weak* form.

These examples demonstrate that there are different outcomes of human-AI interaction that can be represented through the three forms of *relationality*. In the following, we elaborate on the three different forms of the notion *relationality* and highlight how they affect the two different ethical perspectives (*virtue* and *duty*) in the management of human-AI interaction.

With respect to ***relationality of characteristics***, we can observe changes of the capabilities of humans (social entity). For instance, if AI is responsible for the decisions being made in a process, it can lead to a perceived loss of competencies (Mirbabaie et al., 2022, p. 86) and the perceived loss of reputation of the social entity (Mayer et al., 2020). If this perceived loss of reputation by the social entity arises through the human-AI interaction, it is linked to *strong relationality of characteristics*. Suppose the AI gains a higher reputation through the shift of responsibility from humans to AI (representing changes in the material entities' characteristics). In that case, it is linked to *strong relationality of characteristics*).

In contrast to a reputation loss of humans, implementing AI-based systems in an organization can lead individuals to have to "upgrade their understanding of the moral and ethical implications of these new systems" (Watson et al., 2021, p. 469). This leads to increased capabilities of the social entity and thus demonstrates a *strong* form of *relationality of characteristics*. In this case, the individual's ethical behavior (*virtue ethics*) is also influenced by the principle-based behavior of the AI (*duty ethics*). This influence can be either positive or negative, given that the human-AI interaction can either increase or decrease human capabilities depending on how humans "counter the basic need for autonomy" (Pregenzer et al., 2021, p. 12). Ethical considerations regarding the effects of transferring capabilities from humans (social entity) to AI (material entity) arise, such as "What is the trade-off for people? What do they gain from adapting their skills, and what do they lose?"– and thus affects ethical relations of employees' *virtue ethics* (Coombs et al., 2021, p. 5). In comparison, *weak relationality of characteristics* implies that the social entities continue to do "what they are best at while AI shows objective results." (Park et al., 2021, p. 9). Hence, in the *weak* form, there are no effects between the ethical perspectives of social and material entities.

When it comes to ***relationality of agency***, we can observe in the study of Cranefield and Doyle (2022) how its *strong* form can be related to ethics. The authors discovered that teachers interacting with an intelligent writing system while writing reviews of their students' work changed their agency when the system monitored and evaluated their writing (Cranefield and Doyle, 2022). The teachers shifted their focus to improving their own writing quality rather than evaluating their students' writing. Thus, the way individuals change their agency also affects their behavior in interacting with others (Mayer et al., 2020). This shows how the material entity's *duty ethics* curtails the human agency, which leads to changes in the *virtue ethics* of the social entity.

With respect to ***relationality of existence***, the increased autonomy and "fast evolving capabilities" of technology lead to new roles of the AI," such as "new forms of taxing robots [that act] in a way that society once would tax humans" (Coombs et al., 2021, p. 5), representing *strong relationality of existence* of the material entity. However, human-AI interaction may also change the roles of individuals as social entity (also representing a *strong* form of *relationality of existence*). For instance, in the context of data analysis, the human-AI interaction requires new roles for data scientists as tasks switch from manual data classification toward training the AI and "providing auditing process with a ground truth against which the algorithmic outcome [is] compared." (Grønsund and Aanestad, 2020, p. 14). In doing so, the individual as the social entity must evaluate the correctness of the AI (material entity) based on its inherent understanding of correct behavior (*virtue ethics*).

To sum up, (positive or negative) outcomes of human-AI interaction result in different changes for the AI (material entity) and the employee (social entity) – especially in the *strong* subforms of the notion *relationality*. Managed well, these changes in an entity's characteristics, existence, or agency can compensate for the ethical concerns that have arisen through the entanglement or strengthen ethical potentials. Managed poorly, they can enhance ethical issues.

Performativity: How boundaries, limitations, and perceptions of the social subsystem, the social entity, and the material entity affect the social agency and, thus, the ethical management of human-AI interaction.

Performativity is related to boundaries, limitations, and perceptions of the social subsystem, the social entity, and the material entity affecting the agency of the social entity. Depending on which entity (material entity or social entity) or objects (human and non-human objects) of the *social subsystem* create the perceptions about AI, we differentiate between three different forms *performativity*:

i) *External performativity*: the objects from the social subsystem determine the reality from outside (e.g., control mechanisms, guidelines, and norms influence the social entity's perceptions).
ii) *Social performativity*: boundaries of the social entity create perceptions of an own reality (e.g., humans' mental models influence their perceptions).
iii) *Material performativity*: the underlying models of the material entity create perceptions of humans' reality (e.g., data and algorithmic limitations influence the social entity's perceptions)

The following section provides examples based on the literature review to bolster our suggestion to differentiate between the three subforms of *performativity* and explain how the notions shed light on how humans perceive AI based on the boundaries, limitations, and perceptions of their own entity, the material entity, and the objects from the *social subsystem*.

In the case of ***external performativity***, public perceptions of AI may be influenced by external factors, such as media that often induce "dystopia surrounding AI" (Rix and Hess, 2022, p. 8). Employees may perceive working with AI as competing against it rather than cooperating with it (Rix and Hess, 2022, p. 8). How the social entity trusts its organization and perceives the social influence of its colleagues (*social subsystem*) can also influence humans' perceptions regarding AI (Figueroa-Armijos et al., 2022). Depending on how much the individual is influenced by media (*social subsystem*), the *external performativity* can be either *weak* or *strong* – meaning that either there are limiting mechanisms of the social subsystem that affect human perceptions of AI (*strong*) or not (*weak*).

Concerning ***social performativity***, we recognized a *strong* form when, for instance, individuals perceive AI not as a system but as a "machine teammate," as it made individuals "more tolerant of mistakes, to a level similar as to human team members" (Rix and Hess, 2022, p. 7). Not considering AI as a "machine teammate" (*weak social performativity*), by contrast, may tempt humans to "hide behind the output of the machine" and "blame the system, even though … it is not responsible" (Rix and Hess, 2022, p. 9). This is also demonstrated by Jia et al. (2022), who investigated individuals playing games with an app that gives recommendations for the "fastest route to complete the task" (p. 391). They showed that individuals "who self-reported being less sociable tended to attribute less intention to the app when they had lost than when they had won." Negative cognitions of humans, in particular, can lead to perceiving AI as an "inappropriate decision maker" (Bankins et al., 2022, p. 37). Hence, how humans build perceptions of AI can represent either *weak* (the boundaries of humans do not affect its perceptions) or *strong* (the boundaries of humans affect its perceptions) *social performativity*.

With respect to ***material performativity***, the limitations and boundaries of AI as material entity can affect individuals as the social entity working with AI. For instance, Lee et al. (2015, p. 1608) studied Uber and Lyft drivers managed through algorithms and showed that a supply–demand control algorithm's inability to "account for feelings of unfairness" can influence drivers' behavior. Shin and Park (2019, p. 282) identified that algorithms create reality constructions of individuals by trying to "match with users' previous experience," whereby "the more users use algorithms, the narrower their perspectives become" given that "people perceive what they want to in algorithmically recommended contents while try to ignore other viewpoints." Hence, the multiple ways in which the social entity creates its reality based on the AI's functionalities, limitations, and boundaries can result in *weak* (limitations of the AI do not affect humans' perceptions) or *strong* (limitations of the AI affect humans' perceptions) *material performativity*.

These examples demonstrate that there are different ways in which human perceptions of AI can be influenced, which can be represented through the three forms of *performativity*. In the following, we elaborate on the three *strong* forms of the notion *performativity* (*strong material*, *social*, and *external performativity*) and highlight how they affect the two different ethical perspectives (*virtue* and *duty*) in the management of human-AI interaction. *Weak* forms of the notion *performativity* indicate that the social entity's perception of AI is not affected much by others (*social subsystem* or external factors); here, though, we focus on the *strong* form, which actually causes changes in individuals' perceptions regarding AI and, thus, may affect their behavior.

With respect to ***strong external performativity***, we observe that obligatory guidelines (*duty ethics*) that "serve as decision-making support for employees in critical cases" can lead to employees who "do not have to be concerned about whether decisions are ethically justifiable from their point of view, but can rely on the guidelines in making and justifying their decisions" (Mayer et al., 2021, pp. 7–8). Thereby, human's *virtue ethics* are influenced by those guidelines from the *social subsystem*. An individual's *virtue ethics* may also be driven voluntarily by having leaders who "upgrade their understanding of the moral and ethical implications [to] successfully drive organizations forward with innovation in AI" (Watson et al., 2021, p. 469). In this way, leadership from the *social subsystem* influences the ethical ideology of employees (*virtue* ethics) and their perceptions of AI. Hence, the "social influence" of the social subsystem can affect "what individuals deem to be ethical" (Figueroa-Armijos et al., 2022, p. 11). This may also cause individuals to have ambivalent perceptions and expectations regarding AI, depending on whether they interact with an AI agent in their private lives in addition to a professional context (Hinsen et al., 2022). In their private lives, individuals may "do everything that is practical and fun" with AI agents because they possess "full acceptance," whereas AI use in a professional context may be accompanied by concerns about being replaced (Hinsen et al., 2022, pp. 200–201). Hence, *strong external performativity* can have significant effects on an individual's *virtue ethics*, and the context of external influence (e.g., at home or the workplace) is particularly important when considering how to manage human-AI interaction ethically.

When it comes to ***strong social performativity***, we recognize ethical issues regarding *humanity*. For instance, when employees that work with AI claim that "[AI] is not only the system, it is much more," (Rix and Hess, 2022, p. 7) or when"[o]perators who worked directly with the robot regularly characterized their relationship with the robot in collegial or personal terms, referring to the robot as their 'work partner' … or 'friend'" (Sauppé and Mutlu, 2015, p. 3617), we have a *strong* form of *social performativity* that can affect individual *virtue ethics* in a strong way. For instance, overtrust can occur when employees tend to trust the system more than they trust colleagues by ascribing greater capabilities to AI than to their colleagues (*virtue ethics*). "Individuals who are inherently more social tend to view their technological teammates in more social ways as well" and are thus more likely to attribute agency to the AI (Jia et al., 2022, p. 393). To mitigate the negative effects of overtrust, "explanations [about the AI] should be informative, instead of just convincing" (Bansal et al., 2021, p. 13).

With respect to ***strong material performativity***, *responsibility & autonomy* considerations play a crucial role. For example, individuals aware of the limitations of AI may introduce their own control mechanisms to overcome those limitations as a way to preserve their autonomy (*virtue ethics*). Pregenzer et al. (2021), for instance, illustrate that algorithmic-managed Uber drivers use dashcams for defense in case passengers make an accusation against them that would result in the (temporary) deactivation of their accounts ("which is analogous to a suspension in traditional work environments"; p. 7). In this case, the AI's lack of protection abilities as a limitation regarding the AI's *duty ethics* leads to a control mechanism by humans that ensures their autonomy (*virtue ethics*). There are also examples in which humans use the limitations of the AI to shift blame from themselves to the AI, which demonstrates changes in their *virtue ethics*. For instance, "human team members can hide behind the output of the machine teammate and also use it as a scapegoat" by blaming the AI, even though it is not responsible (Rix and Hess, 2022, 9). They thus "ignore" their own ethical ideology (*virtue ethics*) and delegate all ethical considerations to the AI's *duty ethics*.

To sum up, how humans (social entity) build perceptions of AI (material entity) depends on the level of influence through the *social subsystem*, the material entity, or the human itself. When considering the *strong* form of *external performativity* in particular, management practices can directly influence humans' perceptions of AI with respect to possible ethical issues and potential.

Materiality: How interdependencies between objects of the social subsystem and the material entity affect ethical management of human-AI interaction.

While discussions on *materiality* are quite diverse, in this work we refer to *materiality* with respect to how the material entity is arranged based on an artifact's physical and non-physical entities and objects. Materiality does not focus on a specific type of human-AI interaction, as it does not concern *sociomaterial entanglement* and instead affects the interaction indirectly. We distinguish between three subforms of *materiality*:

i) *Physical materiality*: the material entity is arranged based on the (same or other) artifact's physical objects (e.g., documents, hardware, products – representing non-human objects).
ii) *Non-physical materiality*: the material entity is arranged based on the (same or other) artifact's non-physical objects that can be related either to non-human objects (e.g., processes, machine activities, management practices, guidelines) or human objects (e.g., workforces' norms, leadership values).
iii) *Digital materiality*: can be subordinated to *non-physical materiality* and states that the material entity is arranged based on the (same or other) artifact's digital objects (e.g., data, software – representing non-human objects)

The following two sections provide examples from the literature review to bolster our suggestion to differentiate between the three subforms to better explain interdependencies between the objects of the *social subsystem* and the material entity.

Concerning ***physical materiality***, there are interdependencies between the material entity and the artifact's physical objects. For instance, the AI can be influenced by the user interface as an "important tool for scrutinizing the evidence behind the output" (Someh et al., 2022, p. 156). This is a characteristic of *strong physical materiality* between the AI as the material entity and the interface as (non-human) physical object, given that the material entity is interrelated with the same or other artifact's physical objects. How managers drive AI optimization and intent to use AI as a leadership tool (Watson et al., 2021) represents either *weak* (there are no effects) or *strong* (the material entity affects the same or other artifact's physical objects or vice versa) *physical materiality* between the AI as the material entity and the management as a (human) physical object.

Concerning ***non-physical materiality***, there are interdependencies between the material entity and the same or other artifacts' non-physical objects (including digital objects). For instance, data (digital objects) represents a human's discriminating behavior, building on "low-quality input data can lead to biased or unusable outputs" (Asatiani et al., 2020, p. 270). This can evoke "biased standards" of the AI as material entity, leading to "improper procedures" in the AI's decision-making (Zhang et al., 2020, pp. 5888–5889). Thus, AI (material entity) depends on data quality (digital object), which is characterized by *strong digital materiality*. This dependency, though, may also go the other way – for instance, when the data output (digital object) can be improved by the AI (material entity) (Klumpp, 2018). These examples highlight that non-physical objects and the material entity can affect each other in a *weak* (there are no effects) or *strong* (the material entity affects other non-physical objects or vice versa) way.

These interdependencies between the objects of the *social subsystem* and the AI as the material entity confirm the importance of differentiating between several forms of *materiality*. In the following, we elaborate on the two forms of the notion *materiality* (*physical materiality* and *non-physical materiality*, also including *digital materiality*) and highlight how they affect the two different ethical perspectives (*virtue* and *duty*) in the management of human-AI interaction.

We begin with ***weak*** or ***strong physical materiality***. The organizational infrastructure, such as technical infrastructure, represents a physical object that can be interpreted from the *duty ethics* perspective, given that "the organizational infrastructure aligns expectations and goals around AI fairness" (Madaio et al., 2020, p. 7). Moreover, discussing guidelines for the design of AI within the workforce may better "address ethical issues both in the initial design and post-launch stage of AI-system" (Wang et al., 2020, p. 4968). Guidelines embedded in the AI through design represent the *duty ethics* perspective of the artifact's physical objects. Further, empowering ethical guidelines (physical objects) through training as physical objects (comprising the training plan based on *duty ethics,* which might be influenced by the trainer's *virtue ethics*) (Kelley, 2022, p. 880) can increase the explainability of the AI (Asatiani et al., 2020, p. 274). Consequently, "organizations should offer sufficient training" to make employees "feel capable of understanding the new level of technological complexity" and "democratize the process in such a way that it is actually accessible to everyone" (Rix and Hess, 2022, p. 10). The degree of influence between guidelines and training of the *social subsystem* and the increased transparency of the AI are linked to either *weak* or *strong physical materiality*.

Concerning ***weak*** or ***strong non-physical materiality*** in the context of *transparency & accountability*, feedback and explanations (non-physical objects; linked to *strong non-physical materiality*) provided by the workforce as human objects in the *social subsystem* during the use of AI systems (e.g., predicting diagnosis in healthcare) can be used to mitigate ethical issues, for instance, concerning perceived understandability of the AI logic (*duty ethics*) that can occur through this *strong non-physical materiality* (Zhang et al., 2020). In other words, feedback provided by the workforce based on their norms and values (collective *virtue ethics*) influences the AI's logic (*duty ethics*) and can lead to either *strong* or *weak non-physical materiality*. In addition, "discussion through workshops that involve multiple stakeholders" can shed light on the behavior of "Black-Box AI Systems" (Asatiani et al., 2020, p. 259). Discussions among trainees and trainers (*virtue ethics*) are relevant to create "human-consumable explanations," given that "humans ask different explanation-seeking questions" when interacting with AI depending on their background and the task for which they are using AI (Liao et al., 2020, p. 10). This may lead to the establishment of ethical principles in the organization (*duty ethics*) or increase the ethical understanding of the workforce (collective *virtue ethics*). *Virtue ethics* can be further institutionalized if "managers at all levels are key stakeholders in training intelligent systems so that the machines can complement managers' knowledge and work habits" (Kolbjørnsrud et al., 2017, p. 5). Hence, to increase transparency and explainability of the AI, there are several mechanisms (representing non-physical objects of the social subsystem) that can affect the *virtue ethics* of human objects and the *duty ethics* of non-human objects on a collective level.

To sum up, organizations might consider the many opportunities to influence AI's *duty ethics* from the *social subsystem* that indirectly contribute to ethical human-AI interaction. Materiality in this case helps us better explain the interdependencies between objects of the *social subsystem* and the AI (material entity).

Practices: How routines of the material entity or emotions of the social entity affect the social subsystem and, thus, the ethical management of human-AI interaction.

Practices are created through interaction and, thus, represent an outcome of the notion *inseparability*, wherefore they can be linked to all types of human-AI interaction. *Practices* describe real cases and practices that occur through human-AI interaction in daily work. We distinguish between two subforms of practices:

i) *Affective practices*: perceived changes of the material entity within sociomaterial entanglement evoke emotions of human object in the social subsystem (e.g., the workforce).
ii) *Routinized practices*: practices of the material entity within sociomaterial entanglement lead to routinized patterns of non-human objects in the social subsystems (e.g., working operations).

The following two sections provide examples to bolster our suggestion to differentiate between the two subforms of the notion *practices*, given that the differentiation sheds light on how the emotions of individuals as the social entity and routines of the AI as the material entity occur as a result of *sociomaterial entanglement* can affect the *social subsystem*. We also link the subforms of the notion to the types of human-AI interaction.

Beginning with ***affective practices***, we determined that in a *strong* form, it can be linked to *substitution* as one type of human-AI interaction. For instance, the authors of a study of Uber suggested that drivers whose driving routes are dictated by algorithms and who receive unexpectedly low ratings may be frustrated because they know this can lead to account deactivation (Pregenzer et al., 2021). This represents *strong affective practices*, meaning that perceived changes of the material entity evoke strong emotions of the social entity in real cases based on the lack of opportunity to intervene in the algorithmic management through the platform (Lee, 2018; Lee et al., 2015). In a *weak* form, the human-AI interaction does not provoke emotions of the social entity in everyday work. *Strong affective practices* with emotions, especially negative ones, occur mainly when certain mechanisms for interventions are substituted through AI.

Regarding ***routinized practices***, we observed that this subform of *practices* could often be linked to *augmentation* as one type of human-AI interaction. For instance, being managed by algorithms can change the habits of Uber drivers, who begin to explain to their passengers the meaning of their ratings and try to influence those passengers to leave positive ratings (Pregenzer et al., 2021, p. 7). We refer to this as *strong routinized practices*, meaning that practices of the material entity lead to novel routinized patterns of the social entity in everyday work. In a *weak* form, interacting with AI does not lead to routines among the workforce, such as the non-professionalized use of AI-based digital assistants of single employees in organizations with a low AI maturity (Someh et al., 2022). Hence, *strong routinized practices* occur, especially when it comes to *augmentation*.

These examples demonstrate that there are different ways that *practices* may be created in everyday work based on human emotions or routinized behavior. In the following, we elaborate on the two *strong* forms of the notion *practices* (*strong affective practices* and *strong routinized practices*) and highlight how they affect the two different ethical perspectives (*virtue* and *duty*) of the *social subsystem*. As *weak* forms of this notion demonstrate that interacting with AI does not affect the *social subsystem* over the long term, we do not elaborate on the *weak* form.

We recognized that ***strong affective practices*** are often linked to *humanity,* as individuals may feel unease in the impersonal AI-based environment created by "the lack of two-way communication" (Acikgoz et al., 2020, p. 14). As a consequence, human objects of the *social subsystems* try to compensate for the AI's lack of humanity (representing an ethical issue regarding its *duty ethics*) that an individual might perceive during human-AI interaction. Furthermore, social support from other employees as human objects is needed to mitigate that emotional burden. Another example shows that social support can also be a "starting point of worker resistance" (Jiang et al., 2021, p. 412) when "[s]ome drivers try to reassure other drivers by arguing that it is not worth constantly worrying about their ratings" (Jiang et al., 2021, p. 9). Hence, human objects in the *social subsystems* aim to compensate *humanity* for the AI based on their *virtue ethics*.

When analyzing ***strong routinized practices***, we identified that changes in habits and processes could be linked primarily to *privacy & (non–)maleficence* since humans adapt their routines toward non-discriminating behavior (*virtue ethics*) based on the principles according to which the AI is programmed (*duty ethics*). For instance, in the context of assigning passengers to Uber drivers, the algorithm's matching and routing system avoids cherry-picking. This affects the habits of employees, leading to "non-discriminating" behavior of drivers when picking up passengers, since "all rides [are] good for business, in the long run" (Pregenzer et al., 2021, p. 6). However, to follow AI-based principles, it can be helpful to have a "type of checklist that you go through for accuracy," given that organizations "need a good way of incorporating [fairness] as part of the workflow" (Holstein et al., 2019, p. 8).

Another example is that "the AI system is designed to pick up information that employees might accidentally have missed, skipped, or omitted when they processed a customer's loan application," thus minimizing mistakes (Strich et al., 2021, p. 316). However, the AI's equal treatment through its non-discrimination mechanism (*duty ethics*) can nonetheless lead to discrimination against "certain groups of people, based on their social status, origin or place of residence" (Mayer et al., 2020, p. 251), since the AI is not aware of situational exceptions. Thus, *routinized practices,* including the non-discriminating ethical ideology of the employee (*virtue ethics*), are required to compensate for the systematic exclusion through the AI-based system (*duty ethics*).

To sum up, several mandatory principles (*duty ethics*) can help increase individuals' ethical behavior (*virtue ethics*). However, there must be some kind of balance since the AI's *duty ethics* can also drive unethical behavior of the individual. There are different ways *practices* can be created in the *social subsystem* based on human emotions or routinized behavior occurring as a result of *sociomaterial*

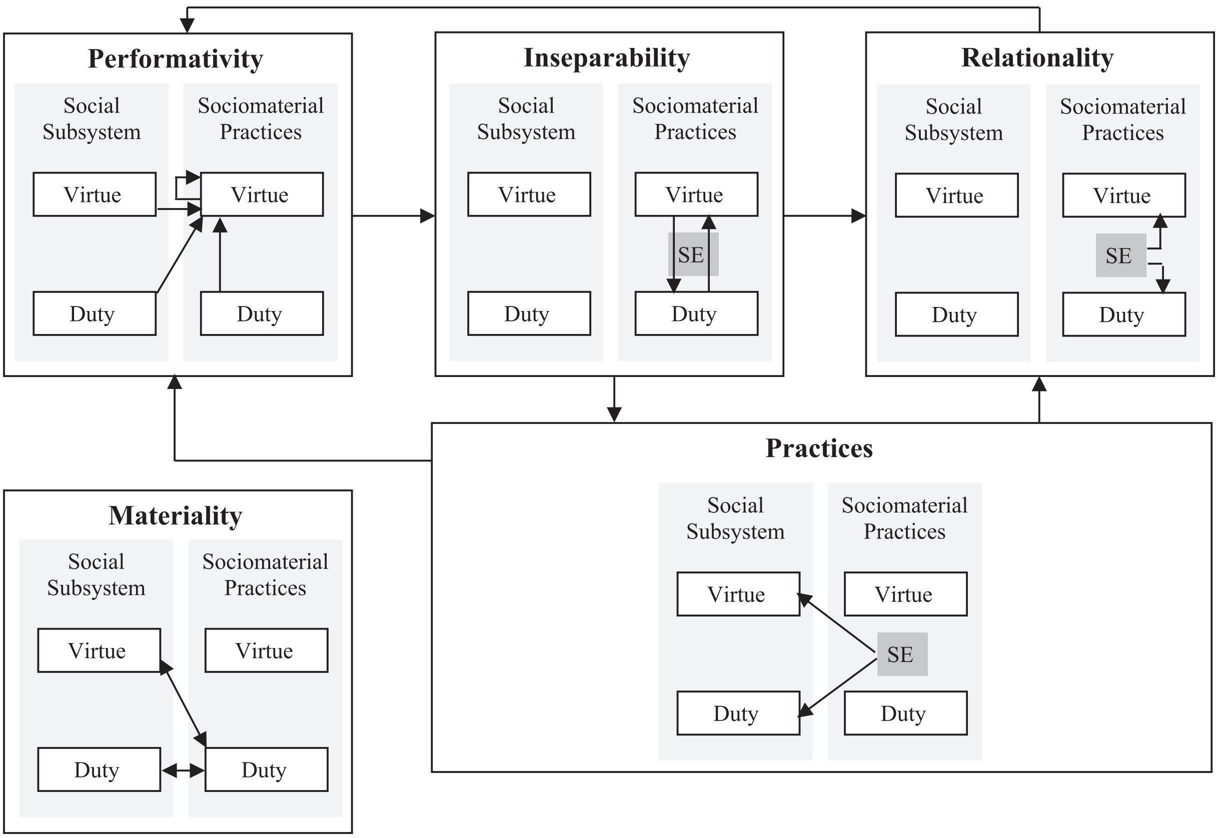

Fig. 4. The Sociomaterial Cycles.

entanglement. Especially when considering *strong practices*, management must be aware of the interdependencies between the ethical intention of an individual's behavior and the *duty ethics* of AI, as both can result in ethical or unethical routines or emotional patterns of the *social subsystem*.

The Sociomaterial Cycles: How the notions of sociomateriality affect ethical management of human-AI interaction through their circular relationship incorporating virtue ethics and duty ethics.

An additional result of analyzing the literature through the different subforms of the five notions of sociomateriality is that we determined that the notions act to some degree in a circular relationship within the sociotechnical system. Fig. 4 is an overview of these relationships affecting both ethical perspectives (*duty* and *virtue*). This can be used to demonstrate the complexity of human-AI interaction and the upstream and downstream processes surrounding the *sociomaterial entanglement:*

In the following, we explain the relationships between the notions illustrated in Fig. 4. In general, we determined two sociomaterial cycles.

We begin with the first sociomaterial cycle and the notion *performativity*, which demonstrates how limitations and boundaries of the material entity, the social entity itself, and the *social subsystem* influence humans' perceptions of AI (social entity) before interacting with the AI as the material entity (*sociomaterial entanglement*). In doing so, *performativity* can outline how an individual's *virtue ethics* builds on human perceptions. Thus, *performativity* describes interdependencies creating AI perceptions of humans that affect human-AI interaction (*sociomaterial entanglement*) as an antecedent. It influences the moment of entanglement that can be illustrated through the notion *inseparability*.

In comparison, *inseparability* focuses on the moment of "imbrication" as the interaction of the social entity and the material entity during *sociomaterial entanglement*. Within *sociomaterial entanglement*, the *duty ethics* of the AI as the material entity and the *virtue ethics* of the human as the social entity influence each other through interaction.

As an outcome of this interaction (and thus, as an outcome of *inseparability*), *relationality* concerns change in the characteristics, agency, or existence of the social entity or material entity. Thus, *relationality* demonstrates how the outcome of human-AI interaction can affect humans' *virtue ethics* and AI's *duty ethics*. Given that these changes in *relationality* can also influence the perceptions of humans as the social entities, *relationality* can affect *performativity*, wherefore we observe the first cycle of influences between *performativity*, *inseparability*, and *relationality*.

In addition, *practices* are created through interaction (and thus represent an outcome of the notion of *inseparability*). Since *practices* lead to routinized patterns of non-human objects (incorporating their *duty ethics*) and emotions of human objects (incorporating their *virtue ethics*) that can be investigated in real everyday work environments, they become part of the *social subsystem* and, thus, influence the perceptions of the social entity as a form of *external performativity*. *Practices* can also change the characteristics of the social entity by provoking emotions that lead to adaptions in the behavior of humans, thus representing *relationality of characteristics* of the social entity. These interdependencies between *inseparability*, *performativity*, *relationality*, and *practices* represent another cycle of influence.

Finally, *materiality* moderates these cycles, as it comprises the several objects of the *social subsystem* (comprising *virtue ethics* and *duty ethics*) that are interdependent with the AI as the material entity (comprising *duty ethics*), which is directly involved within all other notions of sociomateriality through the entanglement with the social entity. Hence, it moderates all other notions of sociomateriality, as it mainly represents management guidelines, practices, and strategies to guide the ethical management of human-AI interaction.

By presenting the interrelationships of the notions of sociomateriality in the form of sociomaterial cycles showing how both ethical perspectives (*duty* and *virtue*) are embedded in those notions, we hope to contribute theoretically to increasing the understanding of the ethical management of human-AI interaction.

Discussion

Research on ethics in human-AI interaction has long been part of the literature (e.g., Abdelaal, 2021, Boada et al., 2021, Giermindl et al., 2022). This research, though, has tended to focus on *duty ethics* (e.g., Chao, 2019, Hamilton et al., 2021, O'Sullivan et al., 2019) – that is, an individual's or technology's "ethical doing" – without considering the influence of moral assumptions and values of an individual's "ethical being" (*virtue ethics*). To investigate the ethical management of human-AI interaction holistically, we systematically analyzed the literature on this topic in the context of both perspectives (*virtue ethics* and *duty ethics*). We applied both perspectives to the five notions of sociomateriality that reflect the complexity of human-AI interaction organizations aim to manage ethically. Where needed, we introduced subforms of the five notions to allow for a more differentiated and reflective analysis. We concur with the suggestions by Berente et al. (2021) regarding the relevance of applying a sociotechnical perspective to guide ethical considerations in managing human-AI interaction.

Theoretical Implications: A research agenda on ethical management of Human-AI interaction

Our review highlights the most significant contributions that research has made toward understanding the ethical management of human-AI interaction. This section extends these contributions by outlining a future IS research agenda comprising four main avenues. The first avenue explains four effects between *virtue ethics* and *duty ethics* occurring through the arrangement of humans, technologies, and human and non-human objects in organizations. The second highlights the contribution of the sociomaterial notions to study the complexity of human-AI interaction. The third demonstrates the relevance of a more human-centered approach when analyzing the ethical management of human-AI interaction in IS research. The fourth avenue proposes using the "Next-Generation Theorizing" of Burton-Jones et al. (2021, p. 303) to enhance the theorizing part of literature reviews according to state-of-the-art requirements for novel complex phenomena, given the changing nature of human-AI interaction that is introducing complexity to strategic management (Benbya et al., 2020). At the end of our discussion of the first and second avenues below, we provide a table that summarizes opportunities for future research.

Avenue 1: The interdependencies between virtue ethics and duty ethics may shape the ethical management of human-AI interaction

Given that one's ethical understanding and intrinsic motivation to behave ethically (*virtue ethics* of the social entity) can be influenced by ethical obligations (*duty ethics* of the material entity or the *social subsystem*), and vice versa, both perspectives shape the ethical management of human-AI interaction. We found four effects when considering both *virtue ethics* and *duty ethics* in a sociotechnical system: *the responsibility attribution effect*; the *control coherence effect*; the *altruistic behavior effect*; and the *behavior adoption effect*. Details are given below.

The *Responsibility Attribution Effect* during sociomaterial practices: humans (social entity) shifting responsibility toward AI (material entity) and, thus, curtailing their ethical being (virtue ethics).

The first effect we outline is the *responsibility attribution effect*, which implies, for example, that humans transfer their responsibility for decisions to the AI-based system to avoid unpleasant situations in which they have to justify certain actions or decisions (Bansal et al., 2021, p. 12). In other words, humans (social entities) transfer their agency of being responsible for certain actions to the material entity's agency (*weak material inseparability*). This happens especially with individuals who have higher levels of sociability, given that "individuals who are inherently more social tend to view their technological teammates in more social ways … [and] are likely to consider the machine as international and agentic." (Jia et al., 2022, p. 394) One striking example of the *responsibility attribution effect* can be seen in the context of a bank where the AI-based loan system takes over decisions to reject or accept loan requests. The bank employee stated, "I think it is quite nice in the case of a loan rejection, because you don't have a long discussion with the customer, the decision is final" (Strich et al., 2021, p. 316).

Other authors refer to an AI acting as a "moral crumple zone taking on responsibility that would have otherwise been assigned to the human" (Hohenstein and Jung, 2020, p. 15). By "feeling that they [decision makers] are no longer responsible for the decisions helps many communicate the decision … more quickly and easily" (Mayer et al., 2020, p. 246). Without AI-based advice, employees "had to develop their own strategies for communicating rejections" (Mayer et al., 2020, p. 246). When AI makes suggestions, employees tend not to question and judge their decisions in everyday work (*strong routinized practices*), and their "ability to reflect critically on their work" decreases (Mayer et al., 2020, p. 249). Employees "no longer have the opportunity to develop reflection skills in work-related decisions because the AI system has taken over the entire decision-making process," so the employees "become indifferent to the outcome of the process." (Mayer et al., 2020, p. 249) For example, with the growing capabilities of LLM-driven conversational agents such as ChatGPT (Mollick, 2022), user may overtrust their output because they view it as close to perfect. This may lead to users simply accepting and adopting the output generated by the AI without reflecting on it and its consequences critically (Dwivedi et al., 2023). In the long term, their "awareness [may be] severely impaired or even nonexistent" (Mayer et al., 2020, p. 249). This demonstrates how an individual's values and abilities of critical thinking and judging (*virtue ethics*) can be affected by attributing responsibility for decision-making and actions from the human to an AI. Still, an AI is guided by guidelines (*duty ethics*) that determine how to handle difficult customer situations (Mayer et al., 2021). As a consequence, we can say that transferring responsibility to an AI agency to make guideline-based decisions affects an individual's ability to judge ethically and critically reflect upon decisions and outcomes (Basu et al., 2021).

The *Control Coherence Effect* between the social subsystem and sociomaterial practices: shifting leadership tasks (duty ethics) from the social subsystem to AI (material entity) can make humans (social entity) behave unethically (virtue ethics).

We call the second effect of interrelations between *virtue ethics* and *duty ethics* the *control coherence effect* – referring in particular to the notion *material inseparability*. This effect occurs when the AI takes over leadership functions, such as monitoring performance outcomes. Given that the AI as the material entity can perform only certain leadership functions by, for instance, collecting data, it often results in controlling the human rather than monitoring the performance of work processes, leading to control coherence. Humans may try to escape the control mechanism of an AI (*duty ethics*) through loopholes (indicating "ill-defined" or nonexistent control mechanism) or blending behavior since they perceive the tight monitoring as unjustified and a loss of their autonomy (*virtue ethics*) (Pregenzer et al., 2021). By having control over the employee, "computers no longer have a subordinate or supporting role but adopt leadership functions" (Hoeddinghaus et al., 2021, p. 7), which can result in employees "feeling exposed" (Rix and Hess, 2022, p. 9) and perceiving a "reduced level of control" (Asatiani et al., 2020, p. 274). This occurs in particular when we observe an interaction between *weak* and *strong material inseparability*, since loopholes provide possibilities for the social entity to have workplace practices that are not mediated by AI as the material entity (*weak material inseparability*) (Pregenzer et al., 2021, p. 11). Yet, when there is a very tight control mechanism of the AI, and when humans are highly dependent on the material entity, the AI inherits the leadership functions from the *social subsystem* through its control mechanism and, thus, gains autonomy over the employee (*strong material inseparability*). Even though individuals may perceive automated leadership decisions as more transparent compared to human decisions, the changes in power and authority may result in individuals "perceive[ing] the traditional form of leadership … as more trustworthy" (Hoeddinghaus et al., 2021, 27).

"Tightening control by introducing even more rules and surveillance may close some loopholes but won't reduce the desire to resist" (Pregenzer et al., 2021, pp. 11–12). As a result, opaque behavior of human-AI interaction (Mayer et al., 2020; Zhang et al., 2020) or "blending behavior" can occur (Pregenzer et al., 2021, p. 10). For instance, in a study of Uber it was outlined that algorithmic-managed drivers who know the algorithm by which they are managed very well may enter "fake inputs" to get preferred passengers, thus optimizing rating scores as a mechanism to escape the pressure to perform due to constant control (Lee et al., 2015, p. 1611). Hence, we suggest that the higher the performance pressure through tight control mechanisms of an AI and the higher the transparency of those control mechanisms (*duty ethics*), the more likely individuals feel exposed and behave in a non-transparent manner by taking advantage of loopholes to escape control (*virtue ethics*). This demonstrates *augmentation* as one type of human-AI interaction (see section 4), in which humans as the social entity try to exploit the algorithm as the material entity with such loopholes to increase their personal benefits (Lee, 2018; Lee et al., 2015; Zhang et al., 2020). In doing so, the human's ethical ideology allows for subordinating morally correct behavior to the advantage of exploiting AI's weakness to escape the control mechanism that they usually allocate to the responsibility of the *social subsystem*. Hence, when AI takes over leadership functions previously done by humans in the *social subsystem*, employees try to escape the control mechanism of the AI's doing (*duty ethics*) by exploiting AI's weaknesses during the interaction (representing the human agency's *virtue ethics*).

The *Altruistic Behavior Effect* during sociomaterial practices: humans (social entity) questioning the ethical doing (duty ethics) of the AI (material entity) based on their ethical being (virtue ethics) can enforce ethical behavior of human-AI interaction.

The third effect of interrelationships between *virtue ethics* and *duty ethics* occurs in particular reference to the notions of sociomateriality *material inseparability* and *routinized practices*. We call it the *altruistic behavior effect*, which implies that humans use their knowledge to question decisions made by AI to achieve greater perceived fairness or ensure security – especially for the benefit of others. In this way, humans as the social entity use their ethical understanding to overturn (*weak material inseparability*) or intervene (*strong material inseparability*) decisions made by the AI as the material agency.

While AI is programmed to treat everyone equally and, thus, follows the non-discrimination principle (*duty ethics*), there are some cases in which humans decide differently based on their inherent ethical understanding (*virtue ethics*). In the context of recruitment, algorithms do not have the ability to, for instance, discern good candidates because they lack "human intuition" by making "judgments based [only] on keywords [making them] not capable of evaluating tones of voices or human interaction" (Lee, 2018, 9). Consequently, humans should intervene (Malik et al., 2021) and rely on their *virtue ethics* to evaluate the AI-based decision (Lee et al., 2015; Pregenzer et al., 2021) and, in doing so, "improve [overall] fairness" in the process of human-AI interaction by combining own fairness perceptions with the fairness being programmed in AI (Holstein et al., 2019, p. 9). Hence, we suggest that when the equal treatment of an AI leads to discrimination, humans as social entities can compensate by using their ethical knowledge for sense-making. When the equal treatment of an AI leads to the potential for non-discrimination, the ethical abilities of humans as social entities can be strengthened.

Fostering humanity in the process of human-AI interaction can support overall fairness – even though humans sometimes "outsmart" the AI to balance insufficient moral behavior of the AI having *augmentation* as an underlying type of human-AI interaction. For instance, at the bank previously mentioned, "some loan consultants were misusing the CleverLoan AI system by trying to outsmart it. They justified these actions as either necessary to overcome the restrictions imposed on their field of activity or important for granting a loan to individuals well-known to them" (Mayer et al., 2020, p. 250). Hence, "outsmarting" is about weighing different moral values against each other (e.g., helping others vs. behaving according to transparency standards), whereby the social entity subordinates the ethical principle to act transparently by discarding the AI-based instruction to increase the greater good (thus fostering beneficence).

The *Behavior Adoption Effect* between the material entity and the social subsystem: managers (social subsystem) adopt ethical decisions (duty ethics) of the AI (material entity).

The fourth effect of interrelationships between *virtue ethics* and *duty ethics* refers, in particular, to the notions of sociomateriality *routinized practices, relationality of agency*, and *material inseparability*. We call it the *behavior adoption effect*, which implies that humans adopt the behavior of an AI by, for instance, also starting to collect sensitive data for performance evaluation, just as the AI-based performance tracker does (Park et al., 2021). In this way, humans as the social entity adapt their ethical understanding of how to evaluate others to the material entity's ethical agency (*strong routinized practices*). For instance, how the AI is programmed can lead to an ignorance of data privacy concerns in a private situation by "collecting and analyzing … sensitive information" (Park et al., 2021, p. 8). One interviewee of Park et al. (2021, p. 8) stated, "Let's say you have a stomachache and stay longer in the restroom. You mean the AI will input that data too? I think that's a serious privacy invasion." Another argues, "Don't you think it's too forceful that it could expose even my private feelings that I want to hide from the manager or firm?" Hence, individuals working with AI or being managed through AI may begin to adopt the ethical or unethical behavior of the AI (based on its *duty ethics*) over the long term as they become used to how AI evaluates performance, for instance. In that way, as their ethical understanding (*virtue ethics*) adjusts to the behavior of the AI, individuals may begin to act similarly.

Ethical issues are further reinforced when the programmed data collection rules of the material entity (*duty ethics*) foster unethical behavior of the social entity when using the information for managerial tasks (*virtue ethics*) as *routinized practices*. If managers also begin to collect such sensitive data and use it in managing employees (*strong routinized practices*), they adapt to the misconduct of the AI (*strong material inseparability*). The "organizational anchoring of ethical AI implementation" (Mayer et al., 2021, p. 9) and the "policies on ethical governance considering socially preferable approaches" (Wang et al., 2020, p. 4968) (*duty ethics*) are important mechanisms of the *social subsystem* that can help avoid humans adapting to misconduct in AI and reduce the ethical issue of the *behavior adoption effect*. Including multiple stakeholders with different ethical understandings, such as domain experts, can also help (Asatiani et al., 2020, p. 274). Hence, we propose that the (un)ethical behavior of AI based on its programmed rules to collect data and make decisions can influence the (un)ethical behavior of managers (and their *virtue ethics* being part of the *social subsystem*) over the long term. Ethical norms and values within the *social subsystem* can moderate those effects.

These effects underline the relevance of understanding how one's inherent ethical ideology is interwoven with the principle-based ethics of the *social subsystem* and the material entity. Hence, future IS research on the ethical management of human-AI interaction should consider the *virtue ethics* perspective more deeply as a way to provide a holistic view of ethics.

Table 1
Toward a Research Agenda for Virtue Ethics in Sociotechnical System Theory.

Effect	Research Questions
Agency Attribution Effect	• How do humans perceive their own responsibility in AI-based decision-making procedures? • How do humans attribute responsibility to the AI, depending on the nature of the task? • Under which circumstances do individuals shift responsibility toward AI?
Control Coherence Effect	• Under which circumstances does human agency dominate material agency and vice versa? • What are the suppressing effects of AI's control mechanism for the outcomes of human-AI interactions (e.g., trust, appreciation, well-being)? • To what extent can the ethical behavior of humans compensate for the non-ethical behavior of AI during augmentative tasks?
Altruistic Behavior Effect	• What patterns of altruistic behavior exist in human-AI interaction? • What positive and negative consequences occur through altruistic behavior in human-AI interaction? • What can or should organizations do to support positive outcomes of altruistic behavior in human-AI interaction?
Behavior Adoption Effect	• How does AI's ethical behavior affect a manager's ethical ideology? • What might be positive and negative consequences when humans begin to adopt AI's ethical view and behavior?

Table 1 is an overview of possible research regarding *virtue ethics* in the context of sociomateriality.

Avenue 2: How the nuanced sociomaterial notions contribute to the analysis of human-AI interaction

The second avenue proposes to apply the sociomaterial notions as a theoretical foundation to make sense of the complexity of ethical considerations of human-AI interaction. With it, we demonstrate the appropriateness of using the lens of sociomateriality – that organizations can be described as an arrangement of humans, technologies, and other objects (such as principles and management practices) that are somehow entangled, and that how they are arranged leads to interdependencies and, thus, to different patterns of actions (Cecez-Kecmanovic et al., 2014; Contractor et al., 2011), such as to ethical or non-ethical behavior. We applied different notions of sociomateriality to provide a deeper and more differentiated understanding of the ethical management of human-AI interaction. The sociomaterial notions contribute to explaining how the different ethical perspectives (*virtue* and *duty*) behave in specific human-AI interactions. Similar to Jones (2014), we identified inconsistencies when analyzing the ethical management of human-AI interaction through the two perspectives of *weak* and *strong* sociomateriality (e.g., the same human-AI interaction can represent *weak* and *strong relationality* depending on whether we reflect the perspective of the material or social entity).

The several ways in which AI interacts with humans are too complex to draw pragmatic differentiations, given the "variety of AI use cases" (Figueroa-Armijos et al., 2022, p. 12) and various tasks for which AI is used that "impact ethical perceptions" (Figueroa-Armijos et al., 2022, p. 1). For example, in the context of autonomous driving, humans are often unaware of AI; a human's cognitive limitations may lead to not recognizing that AI might be part of the car (Hinsen et al., 2022). Instead, humans see the artifact or technology (e.g., application) in which the AI is embedded, such as the vehicle that is parking automatically, thanks to the driving assistant (Hinsen et al., 2022). This example demonstrates *weak material performativity*, since the boundaries of the social entity can lead to perceiving AI as the material entity as nonexistent. Therefore, no perceptions regarding AI's involvement arise when interacting with the technology in which the AI is embedded. But *strong physical materiality* is also illustrated in the example, as the AI changes the vehicle by giving it the agency to behave autonomously.

This example also addresses different ethical aspects. It shows the need for greater transparency in the design of AI-based applications (*duty ethics*) to raise humans' awareness of AI as an important step to making AI more usable in organizations (Someh et al., 2022). In the context of algorithmic-managed leadership decisions in particular, making the use of the algorithm to enhance leaders' decision-making transparent for employees increases their awareness regarding AI involvement, which can lead to increased perceived fairness and trustworthiness of individuals (Hoeddinghaus et al., 2021). This may also result in a shift from *strong* to *weak social performativity*, a shift that can affect an individual's cognitive ability to understand how AI influences the AI-based application as an artifact, given the complexity of the underlying methods and algorithms that can lead to cognitive overload for humans. Hence, training is required to increase an individual's knowledge (Hinsen et al., 2022; Mayer et al., 2020), which points to the *strong physical materiality* of the training since it would not exist without the human-AI interaction.

This example demonstrates that a simple distinction between *weak* and *strong* sociomateriality would not fully explain the underlying complexity of human-AI interaction – especially when "people will have to … understand decisions and consequences from mixed AI/human worlds" (Coombs et al., 2021, p. 4). Therefore, we suggest that the systematic investigation of human-AI interaction through the sociomaterial notions helps increase our understanding of the ethical management of human-AI interaction. By further specifying the single notions, we hope to contribute significantly to the current theoretical discourse. For the same complexity reasons, differentiating solely between different types of human-AI interaction (such as *substitution*, *augmentation*, *assemblage*) as archetypes is not sufficient. As argued in section 4, we have to be aware of how the different tasks are embedded in the process of human-AI interaction. The types of human-AI interaction consider only the moment of sociomaterial entanglement in a specific task (and thus lack considerations of antecedents and consequences of this entanglement), whereas the sociomaterial notions enable a more differentiated analysis of the entire human-AI interaction process (involving the antecedents and consequences of this entanglement), as demonstrated through the sociomaterial cycles.

Table 2 is an overview on future research regarding how the notions of sociomateriality contribute to our understanding of human-AI interaction.

Table 2
Toward a Research Agenda for the Notions of Sociomateriality.

Notion	Research Questions
Materiality	• What is the role of AI ethics in different governance structures of organizations (e.g., codes of conduct, Corporate Social Responsibility guidelines)? (*physical materiality*) • To what extent does AI ethics affect the ethical values (*virtue ethics*) on a collective level (e.g., organizational culture)? (*non-physical materiality*) • To what extent does AI ethics affect the ethics of other technological objects (e.g., design principles of machine ethics)? (*digital materiality*)
Inseparability	• How can we align AI values (*duty ethics*) with human values (*virtue ethics*)? (*social inseparability*) • How can we measure the alignment between human values and AI values? (*material inseparability*)
Relationality	• How can we measure positive outcomes (e.g., satisfaction, self-esteem, well-being) arising through changed existence, characteristics, or agency of individuals as a result of human-AI interaction? (*relationality of individuals' characteristics* and *agency*) • To what extent are positive or negative changes in the existence, characteristics, or agency of individuals a result of human-AI interaction? (*relationality of individuals' characteristics* and *agency*)
Performativity	• How can training influence human agency? (*external performativity*) • How does the transparency of AI affect human agency? (*material performativity*) • How do humans build perceptions of AI? (*social performativity*)
Practices	• How do emotions arising from human-AI interaction affect an individual's *virtue ethics*? (*affective practices*) • How does *virtue ethics* of individuals differ depending on how they are used for interacting with AI? (*routinized practices*)

Avenue 3: How and why a more human-centered approach contributes to the analysis of human-AI interaction

In this avenue, we strive for a more human-centered approach by providing an overview of current research streams concerning human-AI interaction within different disciplines. We demonstrate how our research objective is embedded in the different research fields.

The computer science literature often lacks a human-centered approach. This became clear given how many articles from this discipline we had to exclude because they focus on the design of AI-based technologies, such as articles that explore how to increase fairness and explainability of the AI (e.g., explainable AI (XAI) studies) but often do not explicitly consider humans as the social entity and how their associations and behavior might also affect fairness and explainability of human-AI interaction. When it comes to ethical considerations in particular, a human-centered approach is vital since humans should not be designed out of the technology (Coombs et al., 2021; Ehsan and Riedl, 2020; Liao et al., 2020). For instance, in the context of clinical AI-assisted decision-making, "human-consumable explanations" are crucial to increase trust in the system's output; otherwise, humans could be overwhelmed by the AI explanations (Liao et al., 2020, p. 10). Given that "AI will have a degree of [randomness] and may not be 100 % explainable," explainability should be seen as a social process of conversation and interaction between humans and AI to create "tailor[ed] explanations 'for people with different backgrounds'" (Liao et al., 2020, p. 7). Thus, with respect to the strategic, ethical management of human-AI interaction, "[t]he employee should remain the central element in the workplace, and perceptions of and experiences with AI are fundamental for the process of introducing it to the workplace" (Mirbabaie et al., 2022, p. 87). It is critical, therefore, to ensure a "human-centered" approach for AI given that it "directly impact[s] people's lives and well-being" (Bankins et al., 2022, p. 40).

This is especially relevant given that because of the "non-human and unemotional nature of the AI" (Bankins et al., 2022, p. 15), humans have doubts about their own value, leading to diminished self-esteem. For instance, "the lack of two-way communication in an AI-based interview may signal to the applicant that they will be treated impersonally as an employee (e.g., as 'just a number' or a 'cog in the machine'), if the company chooses not to take the time to have a personal interview" (Acikgoz et al., 2020, p. 412). When AI interacts with humans, the latter can have thoughts such as: "How can we be evaluated by a robot? We are humans and should be treated as valuable individuals" or "It feels like … becoming a part of a gigantic machine" (Park et al., 2021, p. 7). Furthermore, humans sometimes express that their human-AI interactions are "weird, strange, and alien" (Park et al., 2021, p. 7) or that when AI behaves very much like humans, it gives them an "uneasy, creepy feeling" (Rajaobelina et al., 2021, p. 2351). Hence, as algorithms "lack human intuition" (Lee, 2018, p. 9), ethical management needs to preserve key attributes of humanity in human–machine interaction (Coombs et al., 2021) through the *social subsystem*, such as providing emotional support (Kelley, 2022; Park et al., 2021) or feedback loops (Sauppé and Mutlu, 2015). AI's inability to act in a human-intelligent way, such as by being socially and emotionally intelligent, affects the outcomes of human-AI interaction. By integrating mechanism of the *social subsystem* that compensates for the AI's inhumanity, we have a more human-centered approach that contributes to the analysis of human-AI interaction.

The overwhelming majority of the articles we reviewed offer answers to our first subquestion (SQ1), which focuses on how AI interacts with humans. In contrast, far fewer articles provide insights into how humans interact with AI (representing SQ2). There are also articles that focus on the human part (especially in the organizational and management sciences), but they often lack a deeper analysis of ethical management since they are most concerned with investigating AI-related trust and AI attitudes. However, there are interdependencies – for instance, regarding the transfer of ethical knowledge between humans and AI. While the AI's ethical knowledge is expressed in rules and code (e.g., *duty ethics*), the human's ethical knowledge is gained over years of experience and is not easy to express explicitly (e.g., intuition based on *virtue ethics*). Although AI also learns from experience, it lacks "the capabilities to handle situations that are not documented" (Malik et al., 2021, p. 343). Consequently, humans should intervene (Malik et al., 2021) and use their experience to evaluate AI (Lee et al., 2015; Pregenzer et al., 2021) and manage their interactions ethically. Considering only one perspective (either humans interacting with AI or AI interacting with humans) results in ignoring such interdependencies. Hence, we must integrate both perspectives when investigating ethics in human-AI interaction to answer questions such as how to develop ethical AI based on collective social processes of conversation and interaction between humans and AI.

By looking at the intersection, we determined that this is an emerging field of research, especially given that most of the articles in our final sample have been published since 2020. The amount of qualitative work in our final sample further demonstrates that this is an up-and-coming research area.

Avenue 4: How Next-Generation Theorizing contributes to the theorizing part of theoretical literature reviews

Our work further highlights the contributions of using the "Next-Generation Theorizing" of Burton-Jones et al. (2021, p. 303) to enhance the theorizing part of a literature review. More specifically, the *extend strategy* of "Next-Generation Theorizing" enabled us to employ state-of-the-art requirements in the analysis of the changing nature of human-AI interaction as a complex phenomenon that has introduced additional complexity in strategic management (Benbya et al., 2020). We not only consolidated the insights of existing research but also explained the relationships among the constructs in the conceptual framework. In this way, we extend theory on ethics through the lens of sociomateriality to increase our understanding of the ethical management of human-AI interaction.

Based on the four shifts in views of theorizing of Burton-Jones et al. (2021), we derived possible steps that enhance the theorizing part of our review and correspond to the steps for transparency in literature reviews of Templier and Paré (2018). One shift toward "Next-Generation Theorizing" concerns seeing theory as "a lens" to better understand the complexity in this phenomenon and using it as a "starting assumption" (Burton-Jones et al., 2021, p. 306). To address this shift, we embedded sociomateriality directly into the problem formulation as the first step for literature reviews, following Templier and Paré (2018), since the lens of sociomateriality helps in mapping the complexity of the phenomenon. We further refined sociomateriality by using its notions and deriving subforms of each notion to provide a more differentiated analysis in the sixth step of Templier and Paré (2018). In addition, we conducted the first (problem formulation) and second (literature search) step with a holistic view of ethics by involving interdisciplinary research from several domains in the problem formulation and database as well as our keyword search strategy. However, it should be mentioned that the manner in which we consider ethics also represents a theoretical lens that can, depending on which ethical philosophy some refer to, lead to other forms of theories. Still, it would be interesting to investigate what other ethical lenses are used in IS research on human-AI interaction.

Another shift toward "Next-Generation Theorizing" treats theory as "intellectual conversation" and exhorts researchers to focus on the fact that some old theoretical contributions remain relevant while others may change "dramatically" (Burton-Jones et al., 2021, p. 306). We take this shift into account in several steps of Templier and Paré (2018). For instance, we did not restrict the literature search to a certain time range or to one specific research domain. We also conducted "intellectual conversations" in the form of several rounds of discussion when analyzing data and interpreting results, thus adapting our framework.

The shift from "theory as purely desk work" to "theory as reflective engagement" addresses how researchers learn "with and from practitioners to theorize important issues" (Burton-Jones et al., 2021, p. 306). Our analysis is based mainly on articles that report and analyze field research and thus treat human-AI interaction as a changing phenomenon in the real world. We aim to reflect practical cases systematically with the theoretical lens of the notions of sociomateriality and sociotechnical system theory. Thus, we try to foster "reflective engagement" by learning through discussions of the different practical cases while theorizing. This could be extended by embedding reflective methods of empirical approaches into the development of theory to involve different ethical perspectives of ethicists, philosophers, managers, and so on. Future research should investigate how to embed reflective methods (e.g., interviews) in the development of theory.

Finally, the shift "theory as performative" addresses the "need to appreciate how the sociomaterial phenomena we are studying is itself partly a product of theorizing" (Burton-Jones et al., 2021, p. 306). Especially when establishing our framework as the main part of the analysis (sixth step of Templier and Paré, 2018), we recognized patterns to follow the call of Jones (2014) for a more differentiated analysis of sociomateriality (see Avenue 2).

Practical Recommendations: Strategic management of ethical Human-AI interaction

Our review further highlights the relevance of strategic management to fostering ethical human-AI interaction in organizational contexts. In this section, we extend these practical implications by outlining recommendations for three main avenues for future strategies and practices on the ethical management of human-AI interaction. In doing so, we heed the calls of Vial (2019) and of Alsheibani et al. (2020) to enhance our understanding of ethics with respect to the multilevel impact of digital transformation as a way to provide recommendations for leaders in the strategic management of AI.

Our fifth avenue concerns the relevance of bringing ethics to a strategic level in organizations to manage human-AI interaction. Our sixth avenue aims at making employees more literate in their interactions with AI by fostering their (ethical) knowledge for equality in human-AI interaction. Our seventh avenue suggests integrating different stakeholder views, striving for tailored measures, and fostering interdisciplinary research.

Avenue 5: Bringing ethics to a strategic level in managing human-AI interaction in organizations

Ethics is often discussed at either an *operational level* by integrating ethical guidelines into *operative processes*, such as through crash courses on ethics in technological practice projects (Wang et al., 2020) or by implementing ethical checklists (Kelley, 2022; Madaio et al., 2020); or at a *functional level* by providing training and feedbacks in *tactical processes*, such as through regular feedbacks to alert about possible problems and systematic errors (Mayer et al., 2020). Our review highlights the relevance of integrating ethical considerations into corporate and business strategies. Ethics with respect to human-AI interaction requires a greater strategic perspective; as such, ethical checklists and/or guidelines must be internalized and supported by the organizational culture and integrated into the organization's goals and priorities (Grønsund and Aanestad, 2020; Madaio et al., 2020). In addition, a more "value-sensitive design" – with participatory elements being part of the practice and strategies of an organization – is required to develop a holistic understanding of "who" the human is (Ehsan and Riedl, 2020, p. 16). This also contributes to reducing perceived creepiness when interacting with AI-based technologies by integrating human elements, such as touch points (Rajaobelina et al., 2021).

Following the arguments of Kelley (2022), cultural differences put ethical management of human-AI interaction on the business level. Business strategies with respect to policies for AI principles may vary across different countries and may not depend on an overarching corporate strategy. To put ethics on a strategic level, organizations must leverage resources (Madaio et al., 2020) to establish an organizational culture that aims to foster reflection and critical thinking (Mayer et al., 2020). Ethics becoming a part of the business strategy and organizational culture leads to a more objective form of *virtue ethics* at an organizational level, one that is aligned with tailored ethical measures of the AI design based on *duty ethics*. Hence, discussing ethics at a strategic level in organizations will contribute positively to the ethical management of human-AI interaction.

Avenue 6: Making the workforce more literate in human-AI interaction

Our review highlights the relevance of fostering the AI literacy of employees who interact with AI. For instance, Polyviou and Zamani (2022) argue for upskilling the working population with AI skills and enhancing education curricula on AI to foster digital literacy on AI (which is also known as AI literacy). In general, AI literacy can be understood as the skills, capabilities, and competencies that enable individuals to interact with AI effectively (Long and Magerko, 2020). Research on AI literacy highlights the relevance of ethics to successful interaction with AI (e.g., Heyder and Posegga, 2021; Ng et al., 2021), which our review confirms. For instance, a lack of experience can make employees feel as if they are not entitled to direct control over the algorithm and hence do not desire it, which can have ethical consequences (Lee et al., 2015). Furthermore, individuals judge the fairness of algorithmic versus human-made decisions based on their opinion of whether they assume the human or the AI to be more qualified (Lee, 2018). Thus, employees need experiences that can help them achieve a more objective assessment rather than leaving them to make decisions based on gut feelings. Employees require a "sufficient level of metaknowledge" when delegating tasks to AI (Fuegener et al., 2021, p. 680).

A basic understanding of AI and essential information regarding the ethical thinking of AI are also important elements to increase the awareness of ethical AI use in an organization (Mayer et al., 2021, p. 8). Mandatory or voluntary workshops and training events, in particular, can "give employees the knowledge they need to make informed ethical decisions about AI" (Mayer et al., 2021, p. 10). In addition, it is important to install a "culture of learning and experimentation" (Grønsund and Aanestad, 2020, p. 10) to drive "employee's skill and capability renewal" (Watson et al., 2021, p. 471) as well as "adaptive capabilities as humans become involved in new triadic human-AI relationships" (Cranefield and Doyle, 2022, p. 1). While, for instance, "[a]pplications like ChatGPT can be used either as a companion or tutor, to support for example self-regulated learning," they can also lead to individuals minimizing their effort, and thus reduces their independent thinking skills (Dwivedi et al., 2023, p. 9). Hence, the transformative character of AI-based technologies can strengthen but also limit the development of an individual's capabilities. In addition, AI applications "may not be able to produce accurate responses to inputs outside of its training data, or that are highly specific or niche" (Pavlik, 2023, p. 90). Therefore, responsible use of recent AI-based technologies requires sensitivity to the AI's lack of creativity and contextual understanding, the biases in its training data, its limited understanding, its limited ability to personalize instructions, and its dependency on data (Baidoo-Anu and Owusu Ansah, 2023).

Education can help to overcome the "risk of an intelligence divide between humans and machines" (Gupta et al., 2021, p. 12) or "artificial divide" – a serious risk for the strategic management of human-AI – by raising the capability level of employees so they can cooperate successfully with the AI application (Klumpp and Zijm, 2019, p. 270). This is linked to the literature on the AI divide that describes inequalities in the access, usage, and outcomes of human-AI interaction (Carter et al., 2020). Training can strengthen ethical understanding from the bottom up (*virtue ethics*) in addition to top-down approaches such as guidelines and norms (*duty ethics*) (Jiang et al., 2021; Madaio et al., 2020).

Adapting employee skills also requires "new employment schemas," such as new forms of reward systems, to avoid "inequalities and … social fissures" (Coombs et al., 2021, p. 5). In addition to training to upskill workers (Bankins et al., 2022), it is also important to hire employees with the required skillset (Grønsund and Aanestad, 2020). Hence, fostering AI literacy in the workforce can be understood as a prerequisite to the ethical management of human-AI interaction in organizations.

Avenue 7: Integrating different stakeholder views and striving for tailored measures

Our review outlines that ethics is always subject to interpretation and thus stresses the need for an interdisciplinary "ethics team" to interrogate external guidelines (*duty ethics*) with their expertise (*virtue ethics*) (Grønsund and Aanestad, 2020; Mayer et al., 2020; Wang et al., 2020). Moreover, we wish to highlight "the importance of discussing the 'hard decisions' as a group, which may suggest an AI ethics committee may be better than a single AI ethics officer. Regardless, some form of ethics office(r) is important for the effective adoption of AIPs [AI principles]" (Kelley, 2022, p. 881). Emerging AI-based technologies such as ChatGPT offer unique opportunities to incorporate collective ethical considerations in future human-AI interaction outcomes through feedback options based on user input (Radford et al., 2018). Such democratic features that employ the wisdom of the crowd can strengthen the AI model's *duty ethics* based on "collective deliberation and decisions" (Etzioni and Etzioni, 2017, p. 413) represented in user feedback and, as a consequence of further use, strengthen the *virtue ethics* of the user. However, we support the suggestion that ethical measures "would need to differ for different teams or organizations to reflect their organizational culture, goals, and priorities" (Madaio et al., 2020, p. 10). This would mean the involvement of "managers from different levels and geographies in initial experiments with AI" (Kolbjørnsrud et al., 2017, p. 6). Therefore, industry associations, managers, and policymakers should get involved when talking about responsible AI (Gupta et al., 2021). In addition to managers, it is also important that data scientists and "various participants with detailed knowledge on specific segments … join the meetings and contribute their interpretation" (Grønsund and Aanestad, 2020, p. 9). Hence, discussions among different stakeholder groups are relevant "to learn from one another's experiences" (Holstein et al., 2019, p. 8). This form of "cultivating … new expertise" contributes to a greater explainability of the AI, which also increases the value of AI for the organization (Someh et al., 2022, p. 153).

The diversity of roles, sectors, and application areas of AI demands that ethical guidelines be adaptable to different teams and organizations to "reflect their organizational culture, goals, and priorities" (Madaio et al., 2020, p. 10). Ethics, and *virtue ethics* in particular, is partly subjective, and so organizations must "tailor their AI adoption strategies to local and organizational conditions." Since "all technology is local [i]t's about the employee at the desk or in the cubicle and how they use it" (Kolbjørnsrud et al., 2017, p. 5). Hence, to foster ethical management of human-AI interaction, organizations have to establish an ethical ecosystem, activating multi-stakeholder collaboration when governing AI strategies and implementations (Polyviou and Zamani, 2022), as "collective deliberation and decisions" can strengthen *virtue ethics* on a workforce level (Etzioni and Etzioni, 2017, p. 413).

Where do we go from here? Ethical management of human-AI interaction considering recent advances in generative AI

The public and academic discourse on AI-based systems has recently been refueled by advances in technologies based on generative AI (Pavlik, 2023). More specifically, LLMs now serve as the foundation for a variety of services that are accessible to a broad audience within and outside of organizations and that can address a plethora of tasks revolving around processing and generating textual information (Pavlik, 2023). The release of Open AI's LLMs, especially GPT3 and GPT4, has set new standards because of their capability to interpret natural language prompts and generate textual responses that are almost indistinguishable from human writing (Pavlik, 2023). The subsequent integration of these models into ChatGPT, OpenAI's conversational agent, has offered a glimpse into the potential of this branch of generative AI and has served as a role model and foundation for various commercial applications that are likely to have a major impact on knowledge work and human-AI interaction (Baidoo-Anu and Owusu Ansah, 2023; Dwivedi et al., 2023). As a result, this novel generation of generative AI is now widely accessible to a wide range of users (Baidoo-Anu and Owusu Ansah, 2023).

However, the rapid spread of these technologies has amplified some of their inherent issues and fueled the current debate on ethical considerations related to their design and use (Greene et al., forthcoming). Due to the universal applicability of these technologies, it is important to understand how to manage the novel forms of human-AI hybrids from an ethical standpoint, both in theory and practice. In this light, our work can serve as a starting point to guide and structure future research on the ethical management of human-AI hybrids.

In Avenue 1, we discuss the interdependencies between *virtue ethics* and *duty ethics* and how they may shape the ethical management of human-AI interaction. AI-based technologies, including prominent examples such as ChatGPT, are subject to multiple sources of ethical principles. Thus, it is important to consider what happens if they are misaligned or even contradict each other. For instance, *duty ethics*, in the form of organizational norms and guidelines as well as external norms and regulations, set boundaries for developers and guide the development of the system. *Virtue ethics* of individual developers and those of users providing feedback by using and training the system may not always be consistent with the system's *duty ethics*. In the case of ChatGPT, ethical alignment issues become amplified due to the system's widespread use, its integration with various other services, and the public attention it receives (Dwivedi et al., 2023). The interplay between various ethical principles at different stages of developing, training, moderating, and using such technologies remains, for the most part, unexplored. With the growing adoption and success of these proprietary systems, it will be vital to improve our understanding of the ethical influences that shape their behavior to avoid unintended consequences of ethically misaligned models.

This pertains further to Avenues 5 and 7, which address the need to bring ethics to a strategic level of managing human-AI interaction. As Dwivedi et al. (2023, p. 13) stated, "If we want to have any control over this disruption, we need digital leadership and a conversation about acceptable and ethical behavior." Thus, improving our understanding of related phenomena requires research on the strategic management of human-AI interaction, which in turn requires the consideration of stakeholders on all governance layers, such as at a collective (e.g., society, industries, organizations) and individual level (e.g., users). Research on this subject can benefit from a Next-Generation Theorizing approach, which we suggest for analyzing human-AI interaction in times of rapid innovation (see Avenue 4).

In Avenue 2 and Avenue 3, we discuss how the nuanced sociomaterial notions contribute to analyzing human-AI interaction and why a more human-centered approach contributes to the analysis of human-AI interaction. Both avenues offer potential insights into the novel forms of human-AI hybrids that result from the increasing adoption and maturity of generative AI. Due to their ability to interpret natural language prompts, process large amounts of information, and generate compelling textual responses to tasks provided by their users, systems based on generative AI can take on various roles in human-AI hybrid teams: simple ones such as summarizing, editing, and processing textual information; and more complex roles as (co–)producers of content, innovations, and source code (Dwivedi et al., 2023). Understanding the broad spectrum of roles from simple to complex and the consequences for human-AI interaction, as well as their ethical implications, is vital for grasping the technology's impact on workplace practices. Our framework offers a starting point for future research in this direction: the nuanced sociomaterial notions provide a structure for further investigating the ethical management of various types of human-AI hybrids (Avenue 2) while promoting a human-centered approach (Avenue 3).

Avenue 6 calls for making the workforce more literate in human-AI interaction. We observe that it is of the utmost importance to consider the context in which generative AI is applied to figure out which skills and capabilities are required to achieve certain kinds of outcomes that are ethically justifiable. Given that we know little about the skills required to use AI technologies effectively (Long and Magerko, 2020), critically reflect upon their outcomes, and train and develop corresponding "skills, resources, and capabilities to handle generative AI [and] examining biases" (Dwivedi et al., 2023, p. 3), this avenue is particularly promising to advance research on the ethical management of human-AI interaction.

To summarize, we hope that our framework enables the analysis of complex ethical issues that arise with the progress of generative AI, which is leading to novel forms of human-AI interaction. We hope further that our avenues may contribute by integrating a holistic view into the current debate surrounding the use of generative AI applications.

Conclusion and limitations

Our *theoretical review* of the ethical management of human-AI interaction highlights and synthesizes an interdisciplinary and novel body of literature that contributes to our understanding of how we can ethically manage AI interacting with humans (SQ1, concerning how AI's ethical doing based on *duty ethics* might affect human's ethical doing and being) and humans interacting with AI (SQ2, concerning how human's ethical being as *virtue ethics* and doing might affect the AI's ethical doing). Against the backdrop of the current progress of generative AI in gaining increased autonomy (Dwivedi et al., 2023; Mollick, 2022), we hope to provide valuable suggestions with respect to the ethical considerations that arise when managing the changing nature of human-AI interactions. First, we suggest that a more differentiated and systematic analysis, applying the notions of sociomateriality, contributes to explaining the complexity of the different types of human-AI interaction and further clarifies the interdependencies between the different constructs in sociotechnical systems. Second, our work highlights the importance of increased awareness concerning the interplay between *virtue ethics* and *duty ethics* in the context of managing human-AI interaction. We aimed to make this interplay more tangible with the sociomaterial cycles.

With respect to practical implications, our review outlines the importance of bringing ethics to a strategic level in the organization, making the workforce more literate to deal with AI in an ethical way, and integrating different stakeholder views to achieve a holistic view of ethics. Overall, we sincerely hope our findings prove valuable both for future research and for practitioners by enriching the understanding of the ethical management of human-AI interaction.

Our work does have several limitations. First, while we aimed to provide systematicity and transparency in the review process by following the steps of Templier and Paré (2018), the main part of our sixth step (analysis and discussion) was built upon intensive discussion rounds only among the three coders. Following the "Next-Generation Theorizing" of Burton-Jones et al. (2021), in particular, future research could enrich discussion rounds by involving practitioners. Second, our review is restricted to articles that exhibit a certain quality (e.g., by fulfilling specific rankings and peer-reviewed criteria). To include most recent research, future researchers could be less strict in their exclusion and inclusion criteria. Third, our work builds upon the specific perspective of sociomateriality, and thus our results depend on that lens. While this is a limitation, we nevertheless find this lens a valuable tool for investigating the ethical management of human-AI interaction in more depth. Future research that follows a different approach may provide valuable insights into our research objective. Fourth, the objective of understanding ethical management of human-AI interaction called for an approach that favored breadth over depth. Even though we demonstrate relationships of the main constructs, we focused on providing an overall holistic picture. Future research on this topic may zoom in on specific relationships that have emerged from the analysis.

Finally, our *theoretical review* highlights the relevance of rethinking the way we manage human-AI interaction in organizations ethically. To cope with the changing nature of human-AI interaction in terms of increased performance, autonomy, acceptance, and quality, organizations must place ethics on a strategic level. To account for the interdependencies between the inherent ethical values of individuals and the principles or methods AI is designed on, research must consider ethics by reflecting on both perspectives: humans interacting with AI and AI interacting with humans. We hope our framework on the ethical management of human-AI interaction provides valuable guidance for future research and helps inform organizations on how to foster ethics in the strategic management of human-AI interaction. Especially in light of the current debates on generative AI (Mollick, 2022) currently surrounding GPT4 with respect to the unsupervised training of models, close cooperation between users and AI is essential to enable the weighing of different ethics of maxima.

To close out our research, we prompted ChatGPT to consider these very issues in brief, using the prompt "Act as a researcher, create a concluding statement regarding the ethical management of human-AI interaction." Its response was as follows: "*Ensuring ethical management of human-AI interaction is critical for creating a future where humans and machines can work together in harmony. It is our responsibility as researchers to not only develop AI technologies that are reliable and effective but also to consider the ethical implications of their use and ensure that they align with our shared values of fairness, transparency, and accountability*."

The increased capabilities and implementation of AI-based technologies redefine how we work and how we collaborate, and we sincerely hope to provide valuable contributions to better understand the up-and-coming phenomenon.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Notions of weak and strong sociomateriality

In Table A1, we adopt and extend toward our research context the *weak* and *strong* notions of sociomateriality as defined by Jones (2014).

Table A1

Notions of Weak and Strong Sociomateriality, based on Jones (2014).	
Notion: Inseparability	
Inseparability describes the multiple ways social and material entities can be interwoven (Jones, 2014). It focuses on the process of interaction between the social and the material entity and the role each entity plays for the other during the interaction (limited to the imbrication during the process of fulfilling a specific task or goal). The notion suggests that neither entity can function separately. Although Orlikowski and Scott (2008, p. 434) suggest an "inherent inseparability" of the social and material entity, inseparability is not restricted to a specific ontology, as the interrelationship of both entities can vary between *strong* and *weak* (Jones, 2014).	
Strong inseparability describes a "mutual constitution" (Jones, 2014, p. 920). The material and social entities are one unit or, rather, a symbiosis, given that certain tasks or functions do not work without the other entity.	**Weak inseparability** describes a "mutual interdependency" (Jones, 2014, p. 920). The material and social entities operate as separate units in a certain task or function, and they are able to work without the other entity as well in other functions or tasks.
Notion: Relationality	
Relationality describes how sociomaterial entities (social and material entities) are "being produced through relation" (Jones, 2014, 898). Compared to inseparability, *relationality* focuses on the outcome resulting from sociomaterial entanglement. It describes how the entities and their agency, as well as their characteristics, may change because of sociomaterial entanglement. Following Thompson (2012), *relationality* can have different perspectives and varies through sociomaterial entanglement since organizations are not "static entities with determinate[d] properties" (Jones, 2014, p. 898).	
Strong relationality describes that "forms, attributes, and capabilities of the entities emerge only through inter-penetration" (Jones, 2014, p. 920). Without sociomaterial entanglement, there are no changes in the agency or the characteristics of the material or social entity, and no novel entities are created.	**Weak relationality** describes that "forms, attributes, and capabilities of the entities may preexist in any relation" (Jones, 2014, p. 920). Hence, the characteristics of the social or material entity and their agency exist independent of sociomaterial entanglement with the other entity.
Notion: Performativity	
Performativity describes the way how a social entity creates and describes its "reality" (especially linked to the material entity) based on several factors that limit their perceptions (Jones, 2014, p. 899). Different perspectives unfold, limiting rational description. Such boundaries are not "pre-given or fixed, but enacted in practices" (Orlikowski and Scott, 2008, p. 462).	
Strong performativity describes the "enactment of boundaries and relations" (Jones, 2014, p. 920), which means that the existence of the material entity and/or its practices depend on the perception of the social entity. Hence, explanations and limitations of the social entity create the material entity and/or its practices.	**Weak performativity** describes that material entities and practices of material entities exist without being recognized by social entities ("non-human agency"); (Jones, 2014, p. 920). Hence, the reality of the social entity (humans) does not create the material entity, as it exists independent of the perceptions of the social entity.
Notion: Materiality	
Materiality describes how a material entity (tangible or intangible) is arranged based on an artifact's physical or non-physical objects (Faulkner and Runde, 2011; Jones, 2014; Leonardi, 2012). Hence, *materiality* describes how one material entity is arranged around other non-human objects, such as other technologies, principles, or guidelines, of the same artifact or of other artifacts. *Materiality* also comprises human objects of organizational artifacts, such as culture, language, leadership, or the management board, given that *materiality* goes beyond the artifact's physical entity (Barad, 2003; Orlikowski, 2010).	
Strong materiality describes the "materialization of the phenomena" (Jones, 2014, p. 920). The investigated material entity creates new objects within its artifact or changes existing objects across place and time.	**Weak materiality** describes the "persistence of the arrangement of material entities across place and time" (Jones, 2014, p. 920). The investigated material entity does not affect the objects of other artifacts – that is, not changing them across place and time (or making them disappear) – but it may affect the objects of those artifacts.
Notion: Practices	
Practices refer to several forms of mental or physical activities, emotions, knowledge, and so on (Reckwitz, 2002) that can be investigated in everyday work. It refers how perceived changes of the material entity within sociomaterial entanglement affect human and non-human objects of the social subsystem in a real environment (Jones, 2014, p. 914).	
Strong practices are about "embodied, materially, mediated arrays of human activity" (Jones, 2014, p. 920). The material entity mediates physical labor by changing the way of working of the social entity that goes beyond activities by altering human working habits (from an internal standpoint, that is, behavior-related changes).	**Weak practices** focus on "activities and processes" (Jones, 2014, p. 920). The material entity only mediates activities and processes of the entanglement (from an external standpoint, that is, process-related changes).

Appendix B. Quality assessment criteria used for framework development

The assessment criteria we used to develop a framework with desirable quality are based on Schwarz et al. (2007). The way we correspond to those criteria being relevant to our work is given in Table B1.

Table B1
Quality Criteria for Framework Development.

Quality criteria for frameworks aggregated and adapted from Schwarz et al. (2007, pp. 42–43)	Author Considerations
Set up clear guidelines and be aware of the limitations and boundaries of the framework.	• We focused on an *integrating type* of conceptual framework following MacInnis (2011), wherefore the framework does not focus on stating issues or drawing assumptions at this stage in the research process. Only after the second stage the framework was updated and enhanced. • We did not claim for comprehensiveness in the *conceptual stage* (in comparison to the *review stage*).

(*continued on next page*)

Table B1 (*continued*)

Quality criteria for frameworks aggregated and adapted from Schwarz et al. (2007, pp. 42–43)	Author Considerations
Identify relevant research streams and collect different theories.	• We searched in several databases incorporating literature from various research disciplines (see Appendix C) and used relevance-filter options of the databases (if available). • We clarified our main constructs: "human-AI interaction," "ethics," and "strategic management" as our main research streams and restricted the search process to those main constructs (see Appendix M−O). • We involved different perspectives of *weak* and *strong* sociomateriality, and, thus, considered relational as well as substantial ontology (Cecez-Kecmanovic et al., 2014).
Involve all relevant concepts and focus on them.	• We used the references that were most often cited and provided theoretical foundations to identify relevant concepts (Weber, 2012) following a snowball sampling strategy based on backward searches. • We ignored articles that do not use theory for the same purpose as we aim to use in our work, especially those that does not consider strategic or management thoughts (Larsen et al., 2019).
Explain ideas on the relationship between the relevant concepts.	• We explained the interdependencies of our main constructs through the integration of our main constructs in the sociotechnical system of Leonardi (2012). • We further explained interdependencies within the subconstructs of our main constructs following the suggestions of MacKenzie (2003).
Organize specific concepts already studied in several research streams and derive assumptions related to the structure of the framework.	• We organized subconstructs and constructs of the framework by embedding them into the sociotechnical system of Leonardi (2012) in section 4.
Use the framework as a lens to make sense of the synthesized literature.	• We used the arrangement of our main constructs in the sociotechnical system to explain their interdependencies regarding the ethical management of human-AI interaction. This lens is used in the *conceptual stage* as well as in the *review stage* of our research as it is appropriate to think about ethical considerations occurring in the "non-neutral … performativity of the social and technological entanglement" (Cecez-Kecmanovic et al., 2014, p. 826)
Establish a framework that assists to further theory-development.	• We used the framework as an architecture that guides the second stage of our research process following the idea of integrating conceptualizations of MacInnis (2011). • By following the *extend strategy* for theory development of Burton-Jones et al. (2021) we used our framework that reflects well-established theories on ethics and socio-materiality (see section 2) and combined it with the relevant subconstructs that helps us to explain the new phenomenon concerning the changing nature of human-AI interaction (see section 4).

Appendix C. Leading journals and conferences

Leading conferences and journals of the Information Systems (IS) and the Management and Organizational Science research domain are categorized according to the different rankings given in Table C1:

Table C1
Leading Journals and Conferences of the IS Domain and Management and Organization Science Domain.

Leading Outlet Title	Rank	Database
IS-Basket (referring to the Association for Information Systems, 2011): A*/A-Ranking		
MIS Quarterly (Management Information System)	A*	ACM (1977–2019), AIS (1977-present), EBSCO (1977-present), JSTOR (1977–2015), WoS (1980-present)
ISR (Information Systems Research)	A*	ACM (1990–2020), EBSCO (1990-present), JSTOR (1990–2015), WoS (1990-present)
JMIS (Journal of Management Information Systems)	A	ACM (1984–2012), EBSCO (1984-present), JSTOR (1984–2013), WoS (1999-present)
JSIS (Journal of Strategic Information Systems)	A	ACM (1992–2018), EBSCO (2005–2010), SciDir (1991-present), WoS (1995-present)
JIT (Journal of Information Technology)	A	EBSCO (2004–2020), WoS (1993-present)
JAIS (Journal of the Association of Information Systems)	A	ACM (2000), AIS (2000-present), EBSCO (2003-present), WoS (2006-present)
ISJ (Information Systems Journal)	A	ACM (2012–2017), EBSCO (1998-present), WoS (1991-present)
EJIS (European Journal of Information Systems)	A	ACM (1991–2008), EBSCO (1997-present), WoS (1993-present)
TOP IS-Journals (referring to the Lowry et al., 2013): A+/A-Ranking		
MIS Quarterly (Management Information System)	A+	ACM (1977–2019), AIS (1977-present), EBSCO (1977-present), JSTOR (1977–2015), WoS (1980-present)
ISR (Information Systems Research)	A+	ACM (1990–2020), EBSCO (1990-present), JSTOR (1990–2015), WoS (1990-present)

(continued on next page)

Table C1 (*continued*)

Leading Outlet Title	Rank	Database
JMIS (Journal of Management Information Systems)	A+	ACM (1984–2012), EBSCO (1984-present), JSTOR (1984–2013), WoS (1999-present)
EJIS (European Journal of Information Systems)	A	ACM (1991–2008), EBSCO (1997-present), WoS (1993-present)
ISJ (Information Systems Journal)	A	ACM (2012–2017), EBSCO (1998-present), WoS (1991-present)
JAIS (Journal of the Association of Information Systems)	A	ACM (2000-present), AIS (2000-present), EBSCO (2003-present), WoS (2006-present)
JSIS (Journal of Strategic Information Systems)	A	ACM (1992–2018), EBSCO (2005–2010), SciDir (1991-present), WoS (1995-present)
Leading IS-Conferences (referring to the VHB-JOURQUAL3): A-Ranking		
ICIS (International Conference on Information Systems)	A	AIS (1980–2020), EBSCO (1981–2006)
Leading IS-Conferences (referring to the ERA Australian Research Council, Category: Information Systems): A-Ranking		
CSCW (AMC Conference on Computer Supported Cooperative Work)	A	ACM (1986–2020)
CIKM (ACM International Conference on Information and Knowledge Management)	A	ACM (1993–2020), EBSCO (2008–2020), WoS (2012–2019)
SIGIR (ACM International Conference on Research and Development in Information Retrieval)	A	ACM (1971-present), EBSCO (2000)
UIST (ACM Symposium on User Interface Software and Technology)	A	ACM (1988–2020), WoS (2012–2019)
OCIS (AoM Organizational Communication and Information Systems)	A	Not included, check via manual search
ACIS (Australasian Conference on Information Systems)	A	AIS (2001–2020), WoS (2000-present)
HCI (British Computer Society Conference on Human-Computer Interaction)	A	ACM (2007–2015)
CSCL (Computer Supported Collaborative Learning)	A	ACM (1995–2009)
CIDR (Conference on Innovation Data Systems Research)	A	Not included, check via manual search
SCC (IEEE International Conference on Service Computing)	A	ACM (2012-present), IEEE (2004–2020)
ICWS (IEEE International Conference on Web Services)	A	ACM (2004–2005), IEEE (2004–2020), WoS (2004–2019)
ISWC (IEEE International Symposium on Wearable Computing)	A	ACM (2013–2020), IEEE (1997–2012)
VL/HCC (IEEE Symposium on Visual Languages and Human-Centric Computing (was VL))	A	ACM (2009-present), IEEE (2008–2020), WoS (2011–2020)
BPM (International Conference in Business Process Management)	A	Not included, check via manual search
CaiSE (International Conference on Advanced Information Systems Engineering)	A	ACM (2006-present), WoS (1995–2019)
CoopIS (International Conference on Cooperative Information Systems)	A	IEEE (1996–1999), WoS (1997–2019)
DESRIST (International Conference on Design Science Research in Information Systems and Technology)	A	ACM (2009-present)
VHB 2015 (referring to the HARZING-list, Category: General & Strategy): A+/A-Ranking		
Academy of Management Journal	A+	EBSCO (1963-present), JSTOR (1963–2015), WoS (1980-present)
Academy of Management Review	A+	EBSCO (1976-present), JSTOR (1976–2015), WoS (1983-present)
Administrative Science Quarterly	A+	EBSCO (1956-present), JSTOR (1956–2017), WoS (1980-present)
Academy of Management Annals	A	EBSCO (2007-present), WoS (2007-present)
Journal of Management	A	EBSCO (1975-present), SciDir (1993–2004), WoS (1983-present)
Journal of Management Studies	A	EBSCO (1964-present), WoS (1980-present)
Organizational Research Methods	A	EBSCO (1999-present), WoS (1998-present)
Strategic Entrepreneurship Journal	A	EBSCO (2007-present), WoS (2007-present)
Strategic Management Journal	A	EBSCO (1980-present), JSTOR (1980–2015), WoS (1980-present)
VHB 2015 (referring to the -list, Category: Management Information Systems, Knowledge Management): A+/A-Ranking		
Information Systems Research	A+	ACM (1990–2020), EBSCO (1990-present), JSTOR (1990–2015), WoS (1990-present)
MIS Quarterly	A+	ACM (1977–2019), AIS (1977-present), EBSCO (1977-present), JSTOR (1977–2015), WoS (1980-present)
European Journal of Information Systems	A	ACM (1991–2008), EBSCO (1997-present), WoS (1993-present)
Health Care Management Science	A	EBSCO (1998-present), WoS (2007-present)
IIE Transactions	A	EBSCO (1996–2017), WoS (1982–2016)
Information Systems Journal	A	ACM (2012–2017), EBSCO (1998-present), WoS (1991-present)
Journal of Information Technology	A	EBSCO (2004–2020), WoS (1993-present)
Journal of Management Information Systems	A	ACM (1984–2012), EBSCO (1984-present), JSTOR (1984–2013), WoS (1999-present)
Journal of Strategic Information Systems	A	ACM (1992–2018), EBSCO (2005–2010), SciDir (1991-present), WoS (1995-present)
Journal of the Association for Information Systems (AIS)	A	AIS (2000-present), EBSCO (2003-present), WoS (2006-present)
SIAM Journal on Computing	A	ACM (1984–2017), EBSCO (1981-present), WoS (1980-present)

(continued on next page)

Table C1 (*continued*)

Leading Outlet Title	Rank	Database
VHB 2015 (referring to the HARZING-list, Category: Operation Research, Management Sciences, Production & Operations Management): A+/A-Ranking		
Management Science	A+	ACM (1954–2018), EBSCO (1954-present), JSTOR (1954–2015), WoS (1980-present)
Operations Research	A+	ACM (1956–2020), EBSCO (1956-present), JSTOR (1956–2015), WoS (1980-present)
European Journal of Operational Research	A	EBSCO (1977-present), SciDir (1977-present), WoS (1980-present)
INFORMS Journal on Computing	A	ACM (1996–2020), EBSCO (1996-present), WoS (1999-present)
Journal of Operations Management	A	EBSCO (1980-present), SciDir (1980–2018), WoS (1998-present)
Journal of Scheduling	A	ACM (2003–2019), WoS (2001-present)
Journal of the Royal Statistical Society, Series A	A	EBSCO (1974-present), JSTOR (1948–1987)
Journal of the Royal Statistical Society, Series B	A	EBSCO (1998-present), JSTOR (1948–1997)
Manufacturing and Service Operations Management	A	EBSCO (1999-present)
Mathematical Programming	A	ACM (1971–2019), EBSCO (1980-present), WoS (1980-present)
Mathematics of Operations Research	A	ACM (1976–2018), EBSCO (1976-present), WoS (1980-present), JSTOR (1976–2015)
OR Spectrum	A	ACM (1979–2019)
Production and Operations Management	A	EBSCO (2002-present), WoS (1999-present)
Transportation Science	A	EBSCO (1967-present), JSTOR (1967–2015), WoS (1980-present)
VHB 2015 (referring to the HARZING-list, Category: Organization Behavior/Studies, Human Resource Management, Industrial Relations): A+/A-Ranking		
Organization Science	A+	EBSCO (1990-present), JSTOR (1990–2015), WoS (1990-present)
Journal of Organizational Behavior	A	EBSCO (1988-present), JSTOR (1988–2015), WoS (1988-present)
Leadership Quarterly	A	EBSCO (1996-present), SciDir (1990-present), WoS (1994-present)
Organization Studies	A	EBSCO (1980-present), WoS (1981-present)
Organizational Behavior and Human Decision Processes	A	EBSCO (1966-present), SciDir (1985-present), WoS (1985-present)
Personnel Psychology	A	EBSCO (1948-present), WoS (1980-present)
ABDC 2019 (referring to the HARZING-list, Category: General & Strategy): A*/A-Ranking		
Academy of Management Annals	A*	EBSCO (2007-present), WoS (2007-present)
Academy of Management Journal	A*	EBSCO (1963-present), JSTOR (1963–2015), WoS (1980-present)
Academy of Management Review	A*	EBSCO (1976-present), JSTOR (1976–2015), WoS (1983-present)
Administrative Science Quarterly	A*	EBSCO (1956-present), JSTOR (1956–2017), WoS (1980-present)
Journal of Management	A*	EBSCO (1975-present), SciDir (1993–2004), WoS (1983-present)
Journal of Management Studies	A*	EBSCO (1964-present), WoS (1980-present)
Organizational Research Methods	A*	EBSCO (1999-present), WoS (1998-present)
Strategic Management Journal	A*	EBSCO (1980-present), JSTOR (1980–2015), WoS (1980-present)
Academy of Management Discoveries	A	EBSCO (2015-present)
Academy of Management Perspectives	A	EBSCO (1987-present), WoS (2006-present)
Australian Journal of Management	A	EBSCO 1976-present), WoS (2007-present)
British Journal of Management	A	EBSCO (1990-present), WoS (2000-present)
Business & Society	A	EBSCO (1988-present), WoS (2008-present)
Business Strategy & the Environment	A	EBSCO (1995-present), WoS (2009-present)
California Management Review	A	EBSCO (1958-present), WoS (1980-present)
Global Strategy Journal	A	EBSCO (2012-present), WoS (2011-present)
Group Decision and Negotiation	A	EBSCO (1997-present), WoS (1995-present)
Harvard Business Review	A	ACM (1985–1998), EBSCO (1922-present), WoS (1980-present)
International Journal of Conflict Management	A	EBSCO (2006-present), WoS (1994-present)
International Journal of Management Reviews	A	EBSCO (1999-present), WoS (2001-present)
Journal of Forecasting	A	EBSCO (1982-present), WoS (1982-present)
Journal of Management Inquiry	A	EBSCO (1992-present), WoS (1995-present)
Journal of Sport Management	A	EBSCO (1987-present), WoS (1993-present)
Labour History (Australia)	A	EBSCO (1980–2002), WoS (2007–2020)
Long Range Planning	A	EBSCO (1968-present), SciDir (2002-present), WoS (1980-present)

(*continued on next page*)

Table C1 (*continued*)

Leading Outlet Title	Rank	Database
Management and Organization Review	A	EBSCO (2005-present), WoS (2008-present)
Management Learning	A	EBSCO (1970-present), WoS (1994-present)
MIT Sloan Management Review	A	EBSCO (1960-present), WoS (2001–2020)
Organization & Environment	A	EBSCO (1998-present), JSTOR (1997–2017), WoS (2001-present)
Sport Management Review	A	EBSCO (2009-present), WoS (2011-present)
Strategic Entrepreneurship Journal	A	EBSCO (2007-present), WoS (2007-present)
Strategic Organization	A	EBSCO (2003-present), JSTOR (2003–2017), WoS (2007-present)
Strategy Science	A	SciDir (2021-present)
Theory, Culture & Society	A	EBSCO (1982-present), WoS (1991-present)
ABDC 2019 (referring to the HARZING-list, Category: Management Information Systems, Knowledge Management): A*/A-Ranking		
ACM Transactions on Computer Human Interaction	A*	ACM (1994-present), EBSCO (2005-present), WoS (2007-present)
Decision Support Systems	A*	ACM (1989–2018), EBSCO (1995-present), WoS (1991-present), SciDir (1985-present)
European Journal of Information Systems	A*	ACM (1991–2008), EBSCO (1997-present), WoS (1993-present)
Information & Management	A*	ACM (1979–2020), EBSCO (1978-present), WoS (1983-present), SciDir (1977-present)
Information Systems Journal	A*	ACM (2012–2017), EBSCO (1998-present), WoS (1991-present),
Information Systems Research	A*	ACM (1990–2020), EBSCO (1990-present), JSTOR (1990–2015), WoS (1990-present)
Insurance, Mathematics & Economics	A*	EBSCO (1985-present), SciDir (1982-present), WoS (1982-present)
International Journal of Information Management	A*	ACM (1989–2020), EBSCO (1984-present), SciDir (1986-present), WoS (1986-present)
Journal of Information Technology	A*	EBSCO (2004–2020), WoS (1993-present)
Journal of Management Information Systems	A*	ACM (1984–2012), EBSCO (1984-present), JSTOR (1984–2013), WoS (1999-present)
Journal of Strategic Information Systems	A*	ACM (1992–2018), EBSCO (2005–2010), SciDir (1991-present), WoS (1995-present)
Journal of the Am. Soc. For Information Science and Technology	A*	ACM (2001–2013), EBSCO (2001–2013), WoS (2001–2013)
Journal of the Association for Information Systems (AIS)	A*	AIS (2000-present), EBSCO (2003-present), WoS (2006-present)
MIS Quarterly	A*	ACM (1977–2019), AIS (1977-present), EBSCO (1977-present), JSTOR (1977–2015), WoS (1980-present)
AIS Transactions on Human-Computer Interaction	A	AIS (2009-present)
Annales of Applied Probability	A	EBSCO (2020-present)
Behaviour and Information Technology	A	EBSCO (1982-present), WoS (1985-present)
Communications of the ACM	A	ACM (1958-present), EBSCO (1965-present), WoS (1980-present)
Communications of the AIS	A	ACM (1999–2000), AIS (1999-present)
Computers in Human Behaviour	A	SciDir (2021-present)
Data & Knowledge Engineering	A	ACM (1989–2017), EBSCO (1987–2013), WoS (1994-present), SciDir (1985-present)
Electronic Commerce Research	A	ACM (2000–2019), EBSCO (2002-present), WoS (2008-present)
Electronic Markets	A	ACM (2007–2008), WoS (2009-present)
Human Computer Interaction	A	EBSCO (1985-present), WoS (1994-present)
IBM Systems Journal	A	ACM (1962–2008), EBSCO (1966–2008), WoS (1980–2008)
IIE Transactions	A	EBSCO (1996–2017), WoS (1982–2016)
Information and Software Technology	A	ACM (1989–2018), EBSCO (1977-present), SciDir (1987-present), WoS (1987-present)
Information Society (The)	A	ACM (1986–2017), EBSCO (1981-present), WoS (1997-present)
Information Systems Frontiers	A	ACM (1999–2019), EBSCO (2002-present), WoS (2001-present)
Information Technology and People	A	EBSCO (1998-present), WoS (2009-present)
International Journal of Electronic Commerce	A	ACM (1996–2012), EBSCO (2000-present), JSTOR (1996–2013), WoS (2000-present),
Journal of Computer Information Systems	A	EBSCO (2000-present), WoS (1993-present)
Journal of Database Management	A	ACM (1994–2018), EBSCO (1993-present), WoS (2004-present)
Journal of Enterprise Information Management	A	EBSCO (2004-present), WoS (2015-present)
Journal of Global Information Management	A	ACM (1993-present), EBSCO (1993-present), WoS (2005-present)
Journal of Knowledge Management	A	EBSCO (1998-present), WoS (2009-present)

(continued on next page)

Table C1 (*continued*)

Leading Outlet Title	Rank	Database
Journal of Organizational Computing and Electronic Commerce	A	ACM (1996), EBSCO (1991-present), WoS (1999-present)
Journal of Quality Technology	A	ACM (1986-present), EBSCO (1969-present), WoS (1980-present)
Knowledge Based Systems	A	EBSCO (1980–2020), WoS (1991-present)
MIS Quarterly Executive	A	AIS (2002-present), EBSCO (2006-present), WoS (2008-present)
Scandinavian Journal of IS	A	ACM (1991–2006), AIS (1989–2020), EBSCO (2009-present)
ABDC 2019 (referring to the HARZING-list, Category: Operation Research, Management Sciences, Production & Operations Management): A*/A-Ranking		
Annals of Probability	A*	JSTOR (1973–2017), WoS (1980-present)
Annals of Statistics	A*	EBSCO (2003–2020), JSTOR (1973–2017), WoS (1980-present)
Biometrics	A*	EBSCO (1946-present), JSTOR (1947–2015), WoS (1980-present)
Biometrika	A*	EBSCO (1902-present), JSTOR (1901–2015), WoS (1980-present)
Decision Sciences	A*	EBSCO (1970-present), WoS (1984-present)
European Journal of Operational Research	A*	EBSCO (1977-present), SciDir (1977-present), WoS (1980-present)
Journal of Operations Management	A*	EBSCO (1980-present), SciDir (1980–2018), WoS (1998-present)
Journal of the Royal Statistical Society, Series B	A*	EBSCO (1998-present), JSTOR (1948–1997)
Management Science	A*	ACM (1954–2018), EBSCO (1954-present), JSTOR (1954–2015), WoS (1980-present)
Manufacturing and Service Operations Management	A*	EBSCO (1999-present)
Production and Operations Management	A*	EBSCO (2002-present), WoS (1999-present)
Transportation Research Part A: Policy & Practice	A*	EBSCO (1993-present), SciDir (1976-present), WoS (1980-present)
Transportation Research Part B: Methodological	A*	EBSCO (1999-present), SciDir (1979-present), WoS (1980-present)
Transportation Research Part C: Emerging Technologies	A*	EBSCO (2004-present), SciDir (1993-present), WoS (1995-present)
Transportation Research Part E: Logistics and Transportation Review	A*	EBSCO (1997-present), WoS (1997-present)
Advances in Applied Probability	A	WoS (1980-present), JSTOR (1969–2017)
Annals of Operations Research	A	ACM (1988–1993), EBSCO (1984-present), WoS (1991-present)
Applied Statistics: Journal of the Royal Statistical Society Series C	A	EBSCO (1965-present), JSTOR (1952–2016), WoS (1982–1997)
Computers & Industrial Engineering	A	ACM (1986–2020), EBSCO (1985-present), SciDir (1976-present), WoS (1980-present)
Computers & Operations Research	A	EBSCO (1989-present), SciDir (1974-present), WoS (1980-present)
IEEE Transactions on Engineering Management	A	EBSCO (1966-present), WoS (1980-present)
IEEE Transactions on Intelligent Transportation Systems	A	EBSCO (2012–2019), WoS (2000-present)
Industrial Management and Data Systems	A	EBSCO (1981-present), WoS (1994-present)
International Journal of Logistics Management	A	EBSCO (1990-present), WoS (2008-present)
International Journal of Operations & Production Management	A	EBSCO (1980-present), WoS (1994-present)
International Journal of Physical Distribution & Logistics Management	A	EBSCO (1994-present), WoS (2008-present)
International Journal of Production Economics	A	EBSCO (1991-present), SciDir (1991-present), WoS (1991-present)
International Journal of Production Research	A	EBSCO (1961-present), WoS (1980-present)
International Journal of Project Management	A	EBSCO (1997-present), SciDir (1983-present), WoS (2009-present)
Journal of Business Logistics	A	EBSCO (1978-present), WoS (2008-present)
Journal of Multivariate Analysis	A	ACM (1986–2020), EBSCO (1976-present), SciDir (1971-present), WoS (1980-present)
Journal of Optimization Theory & Applications	A	ACM (1967–2019), EBSCO (1975-present), WoS (1980-present)
Journal of Productivity Analysis	A	EBSCO (1989-present), JSTOR (1989–2015), WoS (1994-present)
Journal of Purchasing and Supply Management	A	EBSCO (2003-present), SciDir (2003-present), WoS (2009-present)
Journal of Scheduling	A	ACM (2003–2019), WoS (2001-present)
Journal of Service Management	A	EBSCO (2009-present), WoS (2009-present)
Journal of Supply Chain Management	A	EBSCO (1999-present), WoS (2008-present)
Journal of the Operational Research Society	A	EBSCO (1978-present), JSTOR (1970–1977), WoS (1980-present)
Journal of the Royal Statistical Society, Series A	A	EBSCO (1974-present), JSTOR (1948–1987)
Journal of Transport Geography	A	EBSCO (2002-present), SciDir (1993-present), WoS (2006-present)

(continued on next page)

Table C1 (*continued*)

Leading Outlet Title	Rank	Database
Mathematics of Operations Research	A	ACM (1976–2018), EBSCO (1976-present), JSTOR (1976–2015), WoS (1980-present)
OMEGA – International Journal of Management Science	A	SciDir (1973-present), WoS (1980-present)
Operations Research	A	ACM (1956–2020), EBSCO (1956-present), JSTOR (1956–2015), WoS (1980-present)
Operations Research Letters	A	ACM (1981–2020), EBSCO (2002-present), SciDir (1981-present), WoS (1983-present)
Production Planning & Control	A	EBSCO (1986-present), WoS (1994-present)
Reliability Engineering & System Safety	A	EBSCO (2008–2009), SciDir (1988-present), WoS (1981-present)
Research Technology Management	A	EBSCO (1993-present), JSTOR (1988–2013), WoS (1988-present)
Safety Science	A	EBSCO (1984-present), SciDir (1991-present), WoS (1991-present)
Supply Chain Management: An International Journal	A	WoS (2003-present)
Technometrics	A	ACM (1986–2000), EBSCO (1959–2019), JSTOR (1959–2015), WoS (1980-present)
Theory and Decision	A	WoS (1980-present)
Transport Policy	A	EBSCO (1994-present), SciDir (1993-present), WoS (2005-present)
Transport Reviews	A	EBSCO (1982-present), WoS (1984-present)
Transportation	A	EBSCO (1995-present), WoS (1980-present)
Transportation Research Part D: Transport & Environment	A	EBSCO (1996-present), Sci Dir (1997-present), WoS (1996-present)
Transportation Science	A	EBSCO (1967-present), JSTOR (1967–2015), WoS (1980-present)
ABDC 2019 (referring to the HARZING-list, Category: Organization Behavior/Studies, Human Resource Management, Industrial Relations): A*/A-Ranking		
British Journal of Industrial Relations	A*	EBSCO (1966-present), WoS (1980-present)
Human Relations	A*	EBSCO (1960-present), WoS (1980-present)
Human Resource Management (US)	A*	EBSCO (1972-present), WoS (1985-present)
Industrial and Labor Relations Review	A*	EBSCO (1947-present), JSTOR (1947–2017), WoS (2012-present)
Industrial Relations	A*	EBSCO (1932-present), WoS (1980-present)
Journal of Conflict Resolution	A*	EBSCO (1965-present), JSTOR (1957–2017), WoS (1980-present)
Journal of Human Resources	A*	EBSCO (1966-present), JSTOR (1966–2015), WoS (1980-present)
Journal of Organizational Behavior	A*	EBSCO (1988-present), JSTOR (1988–2015), WoS (1988-present)
Journal of Vocational Behavior	A*	EBSCO (1971-present), SciDir (1971-present), WoS (1980-present)
Leadership Quarterly	A*	EBSCO (1996-present), SciDir (1990-present), WoS (1994-present)
Organization Science	A*	EBSCO (1990-present), JSTOR (1990–2015), WoS (1990-present)
Organization Studies	A*	EBSCO (1980-present), WoS (1981-present)
Organizational Behavior and Human Decision Processes	A*	EBSCO (1966-present), SciDir (1985-present), WoS (1985-present)
Personnel Psychology	A*	EBSCO (1948-present), WoS (1980-present)
Australian Journal of Labour Law	A	Not included, check via manual search
Business Ethics Quarterly	A	EBSCO (1991-present), JSTOR (1991–2015), WoS (2001-present)
European Journal of Industrial Relations	A	EBSCO (1999-present), WoS (1996-present)
European Journal of Work and Organizational Psychology	A	EBSCO (1996-present), WoS (2005-present)
Family Business Review	A	EBSCO (2005-present), WoS (2005-present)
Gender, Work and Organisation	A	EBSCO (1994-present)
Group and Organization Management	A	EBSCO (1976-present), WoS (1992-present)
Human Performance	A	EBSCO (1988-present), WoS (1994-present)
Human Resource Management Journal (UK)	A	EBSCO (1997-present), WoS (2009-present)
Human Resource Management Review	A	EBSCO (1991-present), SciDir (1991-present), WoS (1994-present)
Industrial Relations Journal	A	EBSCO (1970-present)
International Journal of Human Resource Management	A	EBSCO (1990-present), WoS (2000-present)
International Journal of Industrial Organization	A	EBSCO (1983-present), SciDir (1983-present), WoS (1987-present)
International Journal of Intercultural Relations	A	EBSCO (1977-present), SciDir (1977-present), WoS (1983-present)
International Journal of Manpower	A	EBSCO (1980-present), WoS (1994-present)
International Journal of Selection and Assessment	A	EBSCO (1993-present), WoS (1995-present)

(continued on next page)

Table C1 (*continued*)

Leading Outlet Title	Rank	Database
International Negotiation: A Journal of Theory and Practice	A	EBSCO (1996–2017)
Journal of Business and Psychology	A	EBSCO (1986-present), JSTOR (1986–2015), WoS (1994-present)
Journal of Business Ethics *	A	EBSCO (1982-present), JSTOR (1982–2017), WoS (1982-present)
Journal of Industrial Relations	A	EBSCO (1959-present), WoS (2007-present)
Journal of Occupational & Organizational Psychology	A	EBSCO (1991-present), WoS (1992-present)
Labor History	A	EBSCO (1960-present), WoS (1980-present)
New Technology, Work and Employment	A	EBSCO (1997-present), WoS (1994-present)
Organization	A	EBSCO (1994-present), WoS (1995-present)
Organizational Dynamics	A	EBSCO (1972-present), SciDir (1972-present), WoS (1980-present)
Organizational Psychology Review	A	EBSCO (2011-present), WoS (2011-present)
Personnel Review	A	EBSCO (1993-present), WoS (1980-present)
Psychology of Women Quarterly	A	EBSCO (1976-present), WoS (1980-present)
Research in Organizational Behavior	A	EBSCO (1979–2019), SciDir (2000–2019), WoS (1982–2018)
Small Group Research	A	EBSCO (1973-present), WoS (1990-present)
Stress and Health	A	EBSCO (1983-present), WoS (2001-present)
Work & Occupations	A	EBSCO (1974-present), WoS (1982-present)
Work & Stress	A	EBSCO (1987-present), WoS (1987-present)
Work, Employment & Society	A	EBSCO (1987-present), JSTOR (1987–2017), WoS (1991-present)
WoS = Web of Science, SciDir = Science Direct		

Leading conferences and journals of ethics domain were categorized according to different rankings given in Table C2. Since ethical journals and conferences have a lower or no ranking compared to the IS and Management and Organizational Science domain, the rank differs, although they are the leading conferences and journals in the ethics domain.

Table C2
Leading Journals and Conferences of the Ethics Domain.

Leading Outlet Title	**Rank**	**Database**
VHB 2015 (referring to the VHB-JOURQUAL3, Category: Ethics): B/C-Ranking		
Journal of Business Ethics	B	EBSCO (1982-present), WoS (1982-present)
Business Ethics Quarterly (BEQ)	B	EBSCO (2001-present), WoS (2001-present)
Journal of Business, Economics & Ethics	C	Not included, check via manual search
Business Ethics: A European Review	C	EBSCO (1998-present), WoS (2008–2020)
Ethical conferences (referring to the ERA Australian Research Council, Category: Ethics): listed as relevant, but not ranked		
Conference on AI, Ethics and Society (AIES)	–	SciDir (2020-present), WoS (2020-present)
Ethics and Information Technology	–	WoS (2001-present)
WoS = Web of Science, SciDir = Science Direct		

Appendix D. Manual query

The following *five* leading journals and conferences in Table D1 are not covered through the databases from our initial keyword-based search. Hence, they are queried manually by searching for our keywords in the publication titles of the published articles in those nine leading journals or conferences. If an article seems to be relevant but does not include one of the keywords in the publication title, we checked additionally for the keywords in the abstract. Table D1 provides an overview of the articles identified through the manual keyword search. The identified article undergoes the same screening procedure (format screening, quality screening, and relevance screening) as the articles found in the seven databases.

Table D1
Manual Search Results.

Research Domain	Outlet	Hits based on Keywords (Publication Titles)	Hits based on Keywords (Abstract)
IS-domain	Conference on Innovation Data Systems Research (CIDR)	0	0
Management and Organization Science-domain	Australian Journal of Labour Law	0	0
	AoM Organizational Communication and Information Systems	0	0
	International Conference in Business Process Management	0	0
Ethics-domain	Journal for Business, Economics & Ethics	0	1

Appendix E. Keyword-Based search strategy

The keyword field is not available for the databases *ACM Library, AIS Electronic Library,* and *IEEE Computer Society Digital Library* (abstract and title only). We did not use wildcards in the following databases: *AIS Electronic Library, ScienceDirect,* and *JSTOR*. The keyword-based search strategy (with wildcards) is given in Table E1.

Table E1
Keyword-Based Search Strategy.

Keyword-Strings	Human-AI Interaction	AND	Strategic Management	AND	Ethical Principles
keywords	substitution OR augmentation OR assemblage		"corporate strategy" OR "business strategy" OR "functional strategy" OR "planning process*" OR "strategic process*" OR "tactical process*" OR "operational process*"		transparency OR accountability OR privacy OR maleficence OR justice OR fairness OR beneficence OR sustainability OR responsibility OR autonomy
	OR		**OR**		**OR**
Overarching terms = keywords	**"Artificial Intelligence"**		**manag* OR strateg***		**Ethic***

We have the following search string (without wildcards) applied to the title, abstract, and keywords (if available):

("Artificial Intelligence" OR "substitution" OR "augmentation" OR "assemblage") AND ("ethics" OR "transparency" OR "accountability" OR "privacy" OR "maleficence" OR "justice" OR "fairness" OR "beneficence" OR "sustainability" OR "responsibility" OR "autonomy") AND ("management" OR "strategy" OR "corporate strategy" OR "business strategy" OR "functional strategy" OR "planning processes" OR "strategic processes" OR "tactical processes" OR "operational processes").

Appendix F. Format screening

We exclude articles fulfilling the exclusion criteria given in Table F1. In total, the number of articles that have been deleted through filter options given through some of the seven databases is 3,105 articles. The number of articles that have been deleted through manual format screening is 2,379 articles. Together 5,484 articles have been excluded due to format criteria. Please note that the information given in Table F1 only refers to the keyword-search strategy in the databases and the manual search and not to the backward and forward searches.

Table F1
Format Screening.

Exclusion Criteria	Exclusion Description	Number of Eliminated Articles
Non-scientific articles (n = 1,503 articles)	Gray literature, book chapters, book reviews, dissertations, tutorials, bachelor's or master's theses, workshops, posters, letter of editors, special issues, symposiums, talks, magazine, and articles of non-scientific conferences.	Through filter options: **(total: 0 articles)** No filter available Through manual screening: **(total: 1,503 articles)** ACM: 12 articles, AIS: 1,275 articles, EBSCO: 11 articles, IEEE: 135 articles, JSTOR: 1 article, ScienceDirect: 35 articles, Web of Science: 34 articles
Incomplete articles (n = 113 articles)	Short papers or research-in-progress papers, extended abstracts.	Through filter options: **(total: 0 articles)** No filter available Through manual screening: **(total: 113 articles)** ACM: 3 articles, AIS: 83 articles, EBSCO: 4 articles, IEEE: 4

(*continued on next page*)

Table F1 (*continued*)

Exclusion Criteria	Exclusion Description	Number of Eliminated Articles
		articles, JSTOR: 1 article, ScienceDirect: 5 articles, Web of Science: 13 articles
Non-peer-reviewed articles (n = 669 articles)	Articles not being peer-reviewed.	<u>Through filter options:</u> **(total: 417 articles)** ACM: 13 articles, AIS: 4 articles, EBSCO: 392 articles, IEEE: 8 articles <u>Through manual screening:</u> **(total: 252 articles)** ACM: 4 articles, AIS: 28 articles, EBSCO: 7 articles, IEEE: 20 articles, ScienceDirect: 167 articles, Web of Science: 26 articles
Non-English articles (n = 502 articles)	Articles not completely written in English are excluded.	<u>Through filter options:</u> **(total: 385 articles)** EBSCO: 3425 articles, IEEE: 1 article, JSTOR: 37 articles, Web of Science: 2 articles <u>Through manual screening:</u> **(total: 117 articles)** AIS: 15 articles, EBSCO: 2 articles, JSTOR: 1 article, ScienceDirect: 97 articles, Web of Science: 1 article, manual search: 1 article
Non-abstract articles (n = 2,304 articles)	Articles lacking an abstract are excluded due to the mechanism of applying exclusion and inclusion criteria on title, abstract, and keywords.	<u>Through filter options:</u> **(total: 2,303 articles)** ACM: 63 articles, AIS: 45 articles, EBSCO: 525 articles, JSTOR: 49 articles, Web of Science: 1,621 articles <u>Through manual screening:</u> **(total: 1 article)** AIS: 1 article
Review articles (n = 393 articles)	Standalone literature review papers are excluded.	<u>Through filter options:</u> **(total: 0 articles)** No filter available <u>Through manual screening:</u> **(total: 393 articles)** ACM: 1 article, AIS: 222 articles, EBSCO: 30 articles, IEEE: 19 articles, JSTOR: 2 articles, ScienceDirect: 51 articles, Web of Science: 68 articles
Total: 5,484 articles	*<u>Through filter options:</u>* ***3, 105 articles*** *<u>Through manual screening:</u>* ***2,379 articles***	

Appendix G. Quality screening

We only include journal articles published in the Harzing, 2020 journal list (67th ed.; Harzing, 2020) that fulfill the criteria:

- ranking of C or higher in the ABDC (Australian Business Deans Council Journal Rankings List 2019) *OR*
- ranking of C or higher in the VHB (VHB-JOURQUAL3) *OR*
- ranking of 1 or lower in the Ejis & EjisCI (the European Journal of Information Systems List of 2007 of Mingers and Harzing, 2007) *OR*
- ranking of 4 or higher in the UQ (University of Queensland Adjusted ERA Rankings List 2011)

Furthermore, we include all journal articles that are either part of the top21 list of Lowry et al. (2013) or part of the FT 50 journal list (Financial Times, 2016) (in specific, 13 journal outlets not being published in the Harzing, 2020 journal list (67th ed.; Harzing, 2020): e.g., Organization Science).

We include articles from the proceedings of conferences that were ranked C or higher in the ERA conference ranking list (Australian Research Council, 2011) and/or were ranked D or higher in the VHB-conference ranking list (VHB, 2011). We do not apply quality criteria for ethical outlets.

In total, 367 conference articles and 1,491 journal articles have been removed. The eliminated articles (n = 1,858) refer to a total of 1,204 outlets. Thereby, 1,003 journal outlets and 201 conference outlets have been excluded due to the above-mentioned criteria. Table G1 gives an overview of the outlets being excluded.

Table G1
Quality Screening.

Ranking Criteria	Examples of Excluded Outlets	Number of Excluded Outlets
Journal articles not fulfilling the above-mentioned criteria from the Harzing, 2020 journal list (67th ed.; Harzing, 2020) or are not included in the top21 journal list of Lowry et al. (2013) or in the FT 50 journal list (Financial Times, 2016) being excluded.	Advances in Business-Related Scientific Research Journal, European Journal of Communication, International Journal of Clinical Leadership, Journal of Global Information Technology Management, The Journal of Socio-Economics, etc.	**Total: 1,003 articles** <u>Distribution:</u> ACM: 8 articles, AIS: 312 articles, EBSCO: 82 articles, IEEE: 6 articles, JSTOR: 1 article, ScienceDirect: 276 articles, Web of Science: 318 articles
Conference articles having a ranking of at least C or higher in the ERA conference ranking list (Australian Research Council, 2011) or have a	Conference on Emerging Technologies & Factory Automation (ETFA), International Conference on Arts and Culture, International Conference on Multi Agent Systems, International Conference on	**Total: 201 articles** <u>Distribution:</u> ACM: 3 articles, AIS: 88 articles, EBSCO: 16 articles, IEEE: 2 articles, ScienceDirect: 85 articles, Web of Science: 7

(*continued on next page*)

Table G1 (*continued*)

Ranking Criteria	Examples of Excluded Outlets	Number of Excluded Outlets
ranking of at least D or higher in the VHB-conference ranking list (VHB, 2011).	Technical Debt, Wuhan International Conference on e-Business, etc.	

Appendix H. Relevance screening

To be considered relevant, the articles must meet all three criteria given in Table H1. We measure interrater agreement by using Cohen's Kappa following Cohen (1960) to increase review reliability. The intercoder agreement of parallel relevance screening is α = 0.81 (0.05 alpha-level, n = 133). This indicates a strong agreement between coders (McHugh, 2012). In total, 2,121 articles have been eliminated through relevance screening. Please note that the exclusion criteria for relevance screening may apply to more than one criterion.

Table H1
Relevance Screening.

Nr.	Relevance Criteria	Number of Excluded Articles
	The authors must provide their own conceptualization or understanding of ethics regarding human-AI interaction (provide dedicated contributions to our research subject), and any ethical issues or principles should concern the context of human-AI interaction.	**Total: 980 articles** Distribution: ACM: 5 articles, AIS: 295 articles, EBSCO: 55 articles, IEEE: 33 articles, JSTOR: 4 articles, ScienceDirect: 531 articles, Web of Science: 57 articles
1 & 2	Combination of criteria 1 and 2	**Total: 10 articles** Distribution: AIS: 6 articles, EBSCO: 1 article, IEEE: 1 article, ScienceDirect: 2 articles
	Technologies mentioned in the articles must carry out tasks that typically require human cognition, following our underlying definition of AI based on Rai et al. (2019). Hence, we included articles that do not claim the term "AI" specifically, in case they provide ideas on independent and self-learning technologies, such as expert systems, being early forms of intelligent technologies with respect to our thoughts on sociomateriality.	**Total: 806 articles** Distribution: AIS: 464 articles, EBSCO: 56 articles, IEEE: 24 articles, ScienceDirect: 256 articles, Web of Science: 6 articles
2 & 3	Combination of criteria 2 and 3	**Total: 5 article** Distribution: AIS: 1 article, ScienceDirect: 4 articles
	Articles must contain ideas on the strategic management of human-AI interaction in an organizational context having employees as the human part. For instance, the articles must explicitly mention corporate, business, or functional strategies or the associated processes of strategic management of human-AI interaction in organizations or contain ideas that are transferable to the strategic management of human-AI interaction in an organization.	**Total: 269 articles** Distribution: ACM: 2 articles, AIS: 177 articles, EBSCO: 23 articles, IEEE: 5 articles, ScienceDirect: 36 articles, Web of Science: 26 articles
3 & 1	Combination of criteria 3 and 1	**Total: 4 articles** Distribution: AIS: 13 articles, EBSCO: 2 articles, ScienceDirect: 17 articles, Web of Science: 10 articles
1, 2 & 3	Combination of all criteria	**Total: 8 articles** Distribution: ScienceDirect: 6 article, Web of Science: 2 articles

Appendix I. Backward and forward search

Besides conducting backward and forward searches for the references of the remaining 20 articles of the *final sample* based on the keyword search in the databases and the manual search, we also conducted a backward and forward search for the standalone literature review articles and research-in-progress articles, which we excluded through format criteria during the screening process.

Following the references of the 20 articles of the *final sample* from our keyword-based search strategy (keyword sample), we analyzed 177 additional articles (142 articles in the first round, 32 articles in the second round, and 3 articles in the third round). In addition, following the references of the 393 standalone literature reviews (review sample), we analyzed 447 additional articles (244 articles in the first round, 185 articles in the second round, and 18 articles in the third round). Furthermore, following the references of the 113 research-in-progress articles (RIP sample), we analyzed 276 additional articles (246 articles in the first round and 30 articles in the second round). This resulted in 900 articles being analyzed additionally through backward and forward searches.

In total, we added 22 articles to the *final sample* through backward and forward search. Thereupon 7 articles came from the keyword-based search sample, 8 articles from the standalone literature reviews, and 7 articles from the research-in-progress articles. See Table I1 for an overview:

Table I1
Backward and Forward Search.

Search Type	Round	Keyword Sample Nbr. of Articles	Screening	Review Sample Nbr. of Articles	Screening	RIP Sample Nbr. of Articles	Screening
Back-ward	1st round	• 99 articles • Thereof, 9 duplicates	• Excluded format: • 45 articles • Excluded quality: • 9 articles • Excluded relevance: • 31 articles • 5 articles remain	• 179 articles • Thereof, 25 duplicates	• Excluded format: • 39 articles • Excluded quality: • 70 articles • Excluded relevance: • 42 articles • 3 articles remain	• 212 articles • Thereof, 21 duplicates	• Excluded format: • 37 articles • Excluded quality: • 63 articles • Excluded relevance: • 84 articles • 5 articles remain
For-ward	1st round	• 39 articles • Thereof, 4 duplicates	• Excluded format: • 18 articles • Excluded quality: • 6 articles • Excluded relevance: • 11 articles • 0 articles remain	• 65 articles • Thereof, 14 duplicates	• Excluded format: • 20 articles • Excluded quality: • 21 articles • Excluded relevance: • 8 articles • 2 articles remain	• 34 articles • Thereof, 7 duplicates	• Excluded format: • 18 articles • Excluded quality: • 3 articles • Excluded relevance: • 5 articles • 1 article remain
1st round total		142 articles being analyzed (5 articles added to the final sample)		244 articles being analyzed (5 articles added to the final sample)		246 articles being analyzed (6 articles added to the final sample)	
Back-ward	2nd round	• 27 articles • Thereof, 1 duplicate	• Excluded format: • 9 articles • Excluded quality: • 8 articles • Excluded relevance: • 8 articles • 1 article remain	• 77 articles • Thereof, 17 duplicates	• Excluded format: • 4 articles • Excluded quality: • 29 articles • Excluded relevance: • 27 articles • 0 articles remain	• 23 articles • Thereof, 2 duplicates	• Excluded format: • 5 articles • Excluded quality: • 9 articles • Excluded relevance: • 7 articles • 0 articles remain
For-ward	2nd round	• 5 articles • Thereof, 0 duplicates	• Excluded format: • 0 articles • Excluded quality: • 3 articles • Excluded relevance: • 1 article • 1 article remain	• 108 articles • Thereof, 14 duplicates	• Excluded format: • 26 articles • Excluded quality: • 34 articles • Excluded relevance: • 31 articles • 3 articles remain	• 7 articles • Thereof, 1 duplicate	• Excluded format: • 1 article • Excluded quality: • 4 articles • Excluded relevance: • 0 articles • 1 article remain
2nd round total		32 articles being analyzed (2 articles added to the final sample)		185 articles being analyzed (3 articles added to the final sample)		30 articles being analyzed (1 article added to the final sample)	
Back-ward	3rd round	• 3 articles • Thereof, 0 duplicates	• Excluded format: • 0 articles • Excluded quality: • 0 articles • Excluded relevance: • 3 articles • 0 articles remain	• 8 articles • Thereof, 2 duplicates	• Excluded format: • 2 articles • Excluded quality: • 1 article • Excluded relevance: • 3 articles • 0 articles remain	• 4 articles • Thereof, 1 duplicate	• Excluded format: • 0 articles • Excluded quality: • 2 articles • Excluded relevance: • 1 article • 0 articles remain
For-ward	3rd round	-	-	• 10 articles • Thereof, 5 duplicates	• Excluded format: • 1 article	• 3 articles • Thereof, 1 duplicate	• Excluded format: • 1 article

(continued on next page)

Table I1 (*continued*)

Search Type	Round	Keyword Sample Nbr. of Articles	Screening	Review Sample Nbr. of Articles	Screening	RIP Sample Nbr. of Articles	Screening
					• Excluded quality: • 1 article • Excluded relevance: • 3 articles • 0 articles remain		• Excluded quality: • 1 article • Excluded relevance: • 0 articles • 0 articles remain
3rd round total		3 articles being analyzed (0 articles added to the final sample)		18 articles being analyzed (0 articles added to the final sample)		7 articles being analyzed (0 articles added to the final sample)	
Total: 863 analyzed articles		171 analyzed articles (7 articles added to the final sample)		423 analyzed articles (8 articles added to the final sample)		276 analyzed articles (7 articles added to the final sample)	

Appendix J. Final sample

The *final sample* (n = 42) consists of 18 conference articles and 24 journal articles. Thereby, all conference articles have their origin in the IS domain. Most of the journal articles are also from the IS domain (n = 15), whereby there are also 2 journal articles from the Ethics domain and 7 journal articles from the Management and Organization Science domain in the *final sample*. While 20 articles of the *final sample* are from our keyword-based search strategy in the seven databases and manual search, 22 articles are from backward and forward searches. An overview of the *final sample* is given in Table J1:

Table J1
Final Sample.

Nbr.	Autor(s)	Year	Title	Outlet	Source
1	Acikgoz et al.	2020	Justice perceptions of artificial intelligence in selection	International Journal of Selection & Assessment	Keyword sample
2	Asatiani et al.	2020	Challenges of explaining the behavior of black-box AI systems	MIS Quarterly Executive	Keyword sample
3	Bankins et al.	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	Review sample (forward 2nd round)
4	Bansal et al.	2021	Does the whole exceed its parts? The effect of AI explanations on complementary team performance	Conference on Human Factors in Computing Systems	RIP sample (backward 1st round)
5	Basu et al.	2021	Human decision making in AI augmented systems: evidence from the initial coin offering market	Hawaii International Conference on System Sciences (HICCS)	RIP sample (backward 1st round)
6	Cranefield & Doyle	2022	"You have a few wordy spots": role-related impacts of working with an intelligent assistant	European Conference on Information Systems (ECIS)	Keyword sample
7	Coombs et al.	2021	What is it about humanity that we can't give away to intelligent machines? A European perspective	International Journal of Information Management	Review sample (forward 2nd round)
8	Ehsan & Riedl	2022	Requirements for AI-based teammates: a qualitative inquiry in the context of creative workshops	Hawaii International Conference on System Sciences (HICCS)	RIP sample (forward 1st round)
9	Figueroa-Armijos et al.	2022	Ethical perceptions of AI in hiring and organizational trust: the role of performance expectancy and social influence	Journal of Business Ethics	Review sample (forward 2nd round)
10	Fuegener et al.	2021	Cognitive challenges in human-AI collaboration: investigating the path towards productive delegation	Information Systems Research	RIP sample (backward 1st round)
11	Grønsund & Aanestad	2022	Augmenting the algorithm: emerging human-in-the-loop work configurations	The Journal of Strategic Information Systems	Review sample (backward 1st round)
12	Gupta et al.	2021	Role of risks in the development of responsible Artificial Intelligence in the Digital Healthcare Domain	Information Systems Frontiers	Keyword sample
13	Hinsen et al.	2022	How can organizations design purposeful human-AI interactions: a practical perspective from existing use cases and interviews	Hawaii International Conference on System Sciences (HICCS)	Review sample (forward 1st round)
14	Hoeddinghaus et al.	20	The automation of leadership functions: Would people trust decision algorithms?	Computers in Human Behavior	Review sample (backward 1st round)

(continued on next page)

Table J1 (*continued*)

Nbr.	Autor(s)	Year	Title	Outlet	Source
15	Hohenstein & Jung	2020	AI as a moral crumple zone: the effects of AI-mediated communication on attribution and trust	Computers in Human Behavior	Keyword sample
16	Holstein et al.	2019	Improving fairness in machine learning systems: what do industry practitioners need?	Conference on Human Factors in Computing Systems	RIP sample (backward 1st round)
17	Jia et al.	2022	Do we blame it on the machine? Task outcome and agency attribution in human-technology collaboration	Hawaii International Conference on System Sciences (HICCS)	Review sample (forward 1st round)
18	Jiang et al.	2021	Algoactivistic practices in ridesharing - a topic modeling & grounded theory approach	European Conference on Information Systems (ECIS)	Keyword sample
19	Kelley	2022	Employee perceptions of the effective adoption of AI principles	Journal of Business Ethics	Keyword sample
20	Klumpp	2018	Automation and artificial intelligence in business logistics systems: human reactions and collaboration requirements	International Journal of Logistics Research and Applications	Keyword sample (backward 1st round)
21	Klumpp & Zijm	2019	Logistics innovation and social sustainability: how to prevent an Artificial Divide in human–computer interaction	International Journal of Logistics Research and Applications	Keyword sample
22	Kolbjørnsrud et al.	2017	Partnering with AI: how organizations can win over skeptical managers	International Journal of Logistics Research and Applications	Keyword sample (backward 2nd round)
23	Lee	2018	Understanding perception of algorithmic decisions: fairness, trust, and emotion in response to algorithmic management	Big Data & Society	Keyword sample (forward 2nd round)
24	Lee et al.	2015	Working with machines: the impact of algorithmic and data-driven management on human workers	International Conference on Computer-Human Interaction	Keyword sample (backward 1st round)
25	Liao et al.	2020	Questioning the AI: informing design practices for explainable AI user experiences	Conference on Human Factors in Computing Systems	RIP sample (backward 2nd round)
26	Madaio et al.	2020	Co-designing checklists to understand organizational challenges and opportunities around fairness in AI	International Conference on Computer-Human Interaction	Keyword sample (backward 1st round)
27	Malik et al.	2022	Impact of artificial intelligence on employees working in industry 4.0 led organizations	International Journal of Logistics Research and Applications	Keyword sample
28	Mayer et al.	2021	How corporations encourage the implementation of AI ethics	European Conference on Information Systems (ECIS)	Keyword sample
29	Mayer et al.	2020	Unintended consequences of introducing AI systems for decision making	MIS Quarterly Executive	Keyword sample
30	Mirbabaie et al.	2020	The rise of Artificial Intelligence – understanding the AI identity threat at the workplace	Electronic Markets	Keyword sample (backward 1st round)
31	Park et al.	2021	Human-AI interaction in Human Resource Management: understanding why employees resist algorithmic evaluation at workplaces and how to mitigate burdens	Conference on Human Factors in Computing Systems	Keyword sample
32	Pregenzer et al.	2021	Algorithms in the driver's seat: explaining workers' reactions to algorithmic control	European Conference on Information Systems (ECIS)	Keyword sample
33	Polyviou & Zamani,	2022	Are we Nearly There Yet? A Desires & Realities Framework for Europe's AI Strategy	Information System Frontiers	Keyword sample
34	Rajaobelina et al.	2021	Creepiness: Its antecedents and impact on loyalty when interacting with a chatbot	Psychology & Marketing	Review sample (backward 1st round)
35	Rix & Hess	2022	Hello, mate! Insights from the field on leveraging machine teammates in organizations	Pacific Asia Conference on Information Systems (PACIS)	Keyword sample
36	Sauppé & Mutlu	2015	The social impact of a robot co-worker in industrial settings	Conference on Human Factors in Computing Systems	RIP sample (backward 1st round)
37	Shin & Park	2019	Role of fairness, accountability, and transparency in algorithmic affordance	Computers in Human Behavior	Keyword sample (backward 1st round)
38	Someh et al.	2022	Building an Artificial Intelligence explanation capability	MIS Quarterly Executive	Keyword sample
39	Strich et al.	2021	What do I do in a world of Artificial Intelligence? Investigating the impact of substitutive decision-making AI systems on employee's professional role identity	Journal of the Association for Information Systems	Keyword sample
40	Wang et al.	2020	Toward an understanding of responsible Artificial Intelligence practices	Hawaii International Conference on System Sciences (HICSS)	Keyword sample
41	Watson et al.	2021	Will AI ever sit at the C-suite table? The future of senior leadership	Business Horizons	Keyword sample
42	Zhang et al.	2020	Developing fairness rules for talent intelligence management system	Hawaii International Conference on System Sciences (HICSS)	Keyword sample

Appendix K. Meta Information

Table K1 provides an overview of the *meta information* of the *final sample*:

Table K1
Meta Information of the Final Sample.

Meta Information	Codes	Number of Articles	Article Nbr.
Research Discipline	IS Domain	33 articles	2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 23, 24, 25, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 41, 42
	Management and Organization Science Domain	7 articles	1, 20, 21, 22, 27, 34, 40
	Ethics Domain	2 articles	9, 19
Database	ACM	8 articles (thereof, 1 article is also in the Web of Science database)	4, 16, 24,25, 26, 31, 36
	AIS	13 articles (thereof, 2 article are also in other databases)	2, 6, 13, 17, 18, 28, 29, 32, 35, 38, 39, 40, 42
	EBSCO	3 articles	19, 21, 27
	IEEE	1 article (this article is also in the AIS and ScienceDirect database)	32
	JSTOR	0 articles	–
	ScienceDirect	8 articles (thereof, 2 articles are also in other databases)	1, 5, 6, 8, 11, 12, 15, 32, 41
	Web of Science	22 articles (thereof, 1 article is also in the ACM database)	31, 33
	Other databases (articles from the backward and forward search)	11 articles (Elsevier, Sagehub, Springer, SSRN, Wiley Online Library)	3, 7, 9, 10, 14, 20, 22, 23, 30, 34, 37
Publication Type	Journal	24 articles	1, 2, 3, 7, 9, 10, 11, 12, 14, 15, 19, 20, 21, 22, 23, 26, 27, 29, 33, 34, 37, 38, 39, 41
	Conference	18 articles	4, 5, 6, 8, 13, 16, 17, 18, 24, 25, 26, 28, 31, 32, 35, 36, 40, 42
Outlet Ranking	High (e.g., A*, A)	2 articles	11, 40
	Middle (e.g., B, C)	32 articles	2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 27, 28, 29, 30, 32, 33, 34, 35, 37, 38, 40, 41, 42
	Low (e.g., D)	8 articles	1, 4, 16, 24, 25, 26, 31, 36
Paper Type	Empirical	42 articles	1–42
	Non-empirical	0 articles	–
Method Type I	Qualitative	26 articles	2, 6, 7, 8, 11, 13, 18, 19, 20, 21, 23, 24, 26, 27, 28, 29, 31, 32, 33, 35, 36, 38, 39, 40, 41, 42
	Quantitative	10 articles	1, 5, 9, 10, 12, 14, 15, 17, 34, 37
	Mixed-Method	6 articles	3, 4, 16, 22, 25, 30
Method Type II	Online Survey	9 articles	3, 9, 12, 15, 16, 22, 23, 30, 34
	Interviews	14 articles	6, 7, 13, 18, 19, 24, 25, 26, 27, 31, 35, 36, 37, 41, 42
	Case Study	11 articles	2, 5, 8, 11, 20, 21, 28, 29, 38, 39, 40
	Experiment	5 articles	1, 4, 10, 14, 17
	Delphi Study	1 article	42
	Social Media / Text Analysis	2 articles	32, 33
Analysis Type	Qualitative Analysis	20 articles	6, 7, 13, 19, 20, 21, 24, 25, 26, 27, 28, 29, 31, 33, 35, 38, 39, 40, 41, 42
	Descriptive Statistics	9 articles	1, 9, 10, 12, 14, 15, 17, 34, 37
	Hybrid Analysis Forms	13 articles	2, 3, 4, 5, 8, 11, 16, 18, 22, 23, 30, 32, 36
Sample Type	Students	2 articles	1 (also an employee sample), 17
	Employees/Workers	18 articles	1 (also a student sample), 2, 4, 9, 10, 12, 18, 19 20, 21, 23, 24, 26, 27, 28, 31, 32, 34
	Managers	6 articles	11, 22, 36, 38, 39, 41
	Researchers	1 article	7
	Experts	8 articles	5, 6, 13, 25, 28, 29, 34, 42
	No Sample (or not claimed)	2 articles	16, 40
	Random Sample (citizens, users)	5 articles	3, 8, 14, 15, 33, 37
Sample Size	<100	22 articles	2, 6, 7, 8, 13, 17, 19, 20, 21, 24, 25, 26, 27, 28, 29, 31, 35, 36, 39, 41, 42
	100–299	5 articles	9, 12, 15, 18, 23
	299–599	7 articles	1, 14, 16, 30, 33, 34, 37
	600–1000	3 articles	3, 10, 32
	>1000	3 articles	4, 5, 22
	No sample (or not claimed)	2 articles	16, 40
Result Type	Propositions / Future Research Directions	2 articles	15, 16

(*continued on next page*)

Table K1 (*continued*)

Meta Information	Codes	Number of Articles	Article Nbr.
	Framework / Classification / Summary	10 articles	2, 7, 8, 11, 18, 23, 27, 33, 39, 42
	Hypotheses-Validation	9 articles	1, 3, 4, 5, 9, 10, 30, 32
	Theory-Development	7 Articles	6, 14, 17, 20, 21, 32, 35
	(Practical/Design) Guidelines	13 articles	13, 19, 22, 24, 25, 26, 28, 29, 31, 36, 38, 40, 41
Contribution Type	(rather) theoretical	14 articles	5, 6, 7, 8, 10, 14, 15, 16, 18, 27, 30, 32, 34, 35
	(rather) practical	12 articles	2, 13, 19, 22, 28, 29, 31, 33, 36, 38, 40, 41
	Both	15 articles	1, 3, 9, 11, 12, 17, 20, 21, 23, 24, 25, 26, 37, 39, 42
	Not claimed	1 article	4
Use of Theory	Exploratory	21 articles	2, 8, 11, 13, 15, 16, 17, 23, 24, 25, 26, 31, 32, 35, 36, 37, 38, 39, 40, 41, 42
	Deductive	12 articles	1, 5, 9, 10, 12, 14, 19, 20, 21, 27, 30, 31
	Inductive	7 articles	6, 7, 18, 22, 28, 29, 33
	Mixed	2 articles	3, 4

Appendix L. Coding schema

For deductive coding, we follow the steps of Mayring (2014, p. 95): 1. Definition of the research question and the theoretical background 2. Definition of the category system (main categories and subcategories, 3. Definition of coding guidelines, 4. A material run-through, preliminary coding, coding rules, 5. Revision of the categories and coding guidelines after 10–50 % of the material, 6. Final working through the material, and 7. Analysis, category frequencies, and contingencies interpretation. Whereas the first step is already done before (see sections 1, 2, and 4), the remaining steps are part of section 5. The deductive coding scheme is mainly based on section 2 (theoretical lens of sociomateriality and sociotechnical system theory) and section 4 (theoretical background on the main constructs). Table L1 provides a snapshot of the deductive coding scheme.

Table L1
Snapshot of the Deductive Coding Scheme.

Category	Definition	Sub-category	Definition	Coding Rule	Example
Inseparability	Inseparability describes the multiple ways social and material entities can be interwoven (Jones, 2014). It focuses on the interaction process between the social and the material entity and the role each entity plays for the other during the interaction (limited to the imbrication during the process of fulfilling a specific task or goal). The notion suggests that neither entity can function separately. Although Orlikowski and Scott (2008, p. 434) suggest an “inherent inseparability” of the social and material entity, inseparability is not restricted to a specific ontology, as the interrelationship of both entities can vary between strong and weak (Jones, 2014).	Strong inseparability	Strong inseparability describes a “mutual constitution” (Jones, 2014, p. 920). The material and social entities are one unit or a symbiosis, given that certain tasks or functions do not work without the other entity.	• The text must contain at least one of the criteria: • Interdependencies between both entities (certain tasks or functions do not work without the other entity) • Both entities shape each other or form a unit	“Drivers need to turn on their ridesharing app to be able to receive and execute their work” (Lee et al., 2015, p. 1604)OR“Alder Play [an AI-based APP used in hospitals] enables healthcare professionals to have access to medical records of patients who are eligible for NHS treatment. Patients and their families would be able to obtain their medical records online. This could largely improve transparency in the clinical processes, thereby enhancing the quality of health services and strengthening the patient engagement.” (Wang et al., 2020, p. 4965)

Appendix M. Types of Human-AI interaction

The three types of human-AI interaction following Rai et al. (2019) are given in Table M1:

Table M1
Types of Human-AI Interaction.

Type	Characteristics	Similar terms
Substitution	• AI substitutes human agents in performing a task (Rai et al., 2019) • Replacing the tasks people used to do (Norman, 2017) • Keep humans out of the equation to allow more comprehensive, rational, and efficient processing (Davenport and Kirby, 2016, p. 26) • "automation involves the substitution of machines for labor and leads to the displacement of workers from the tasks that are being automated" (Acemoglu, D., Restrepo, P., 2018, p. 5)	• Task elimination (Norman, 2017) • Task automation (Raisch and Krakowski, 2020) • Human automation (Janssen et al., 2019) • Human-out-of-the-loop (Shahriari et al., 2016)
Augmentation	• The main focus is to balance each other's weaknesses and learn from each other by using the complementary strengths of human intelligence and AI (Daily et al., 2017; Kamar, 2016; Raisamo et al., 2019) • *Machine augmentation*: bringing human intelligence in the AI to design, complement, and evaluate the capabilities of AI (Mnih et al., 2015) or by having the human as a teacher bringing the system to learn (Dellermann et al., 2019b) • *Human augmentation*: AI helping humans expand their abilities by extending physical capabilities or interacting with individuals more effectively (Wilson and Daugherty, 2018, p. 214)	• Human-in-the-loop (Daily et al., 2017; Raisamo et al., 2019; Wilson and Daugherty, 2018) • *Human augmentation* (Raisamo et al., 2019) • Augmented Human Intelligence (Dellermann et al., 2019b) • Augmented Machine Intelligence (Dellermann et al., 2019b)
Hybrid Intelligence	• "systems that have the ability to accomplish complex goals by combining human and artificial intelligence to collectively achieve superior results" (Dellermann et al., 2019a, p. 1) • "AI and human agents are assembled contextually and temporally … in assemblage the human and AI agents work as an integrated unit" (Rai et al., 2019, p. iv) • Distribution of roles in hybrid intelligence: augmented human intelligence and augmented machine intelligence (Dellermann et al., 2019b, p. 640) • "The performance of Hybrid Intelligence systems can, thus, not only be measured by the superior outcome of the whole sociotechnical system alone, but the learning … of human and machine agents" (Dellermann et al., 2019b, pp. 640–641) • "It requires deeper reasoning capabilities for machines to make decisions not only about how they are accomplishing their tasks, but also about how they can support their teammates towards the success of the collaborative activity" (Kamar, 2016, p. 4073)	• Assemblage (Rai et al., 2019) • Combining augmented human and augmented machine intelligence (Dellermann et al., 2019b)

Appendix N. Ethical principles

In 2004, the politically driven attempt by the Organization for Economic Co-operation and Development (OECD), launched in 2004, to concretize ethics for businesses centered on four widely applicably principles: fairness, transparency, accountability, and responsibility (Rezaee, 2008, p. 35). It should be noted that we chose our own classification based on the most common principles of business ethics we found in business ethics literature, which are given in Table N1:

Table N1
Principles of Business Ethics.

Ethical principles	Characteristics	Authors
Transparency & understandability (distinctiveness, visibility, accuracy)	• Understandable *management processes* (Rezaee, 2008) • Enhancing accountability through assurance (Dando and Swift, 2003)	e.g., Dando and Swift, 2003; Rezaee, 2008
Security & (non-)maleficence (human/civil/employment rights, no-harassment, security, safety, political contributions, compliance, non-discrimination, no-risk)	• Avoid conflicts of interest (Payne and Wayland, 1999) • Manage personal biases (Payne and Wayland, 1999) • Comply with law and human rights (Goodpaster, 1991)	e.g., Dando and Swift, 2003; Goodpaster, 1991; Payne and Wayland, 1999; Winstanley et al., 1996
Fairness & justice (consistency, consensus, diversity, dignity, respect, inclusiveness, non-discrimination)	• Treat all people with equal respect (Hasnas, 2013) • Promote fairness and justice for all stakeholders (Rezaee, 2008)	e.g., Hasnas, 2013; Payne and Wayland, 1999; Rezaee, 2008; Winstanley et al., 1996
Sustainability & credibility (relevance, well-being, innovation)	• Enforce long-term impact of management practices (Goodpaster, 1991)	e.g., Goodpaster, 1991; Hasnas, 2013; Winstanley et al., 1996

(continued on next page)

Table N1 (*continued*)

Ethical principles	Characteristics	Authors
Responsibility & leadership (authority, legitimacy, agreement, validity, obligations, credibility, control, deregulation)	• Enhance human well-being (Hasnas, 2013) • Strengthen social, ethical, and environmental credibility (Dando and Swift, 2003) • Balance between decision-making power of all organizational stakeholders (Hasnas, 2013; Payne and Wayland, 1999) • Ensure that people act with awareness (Hasnas, 2013) • Increase organizational profits so long as there are no negative consequences for organizational stakeholders (Greenwood, 2002)	e.g., Hasnas, 2013; Greenwood, 2002; Payne and Wayland, 1999; Schwartz and Saiia, 2012; Winstanley et al., 1996

Several authors aggregated and categorized ethical principles in the context of AI. For instance, Leslie (2019) identified key attributes of AI ethics, such as fairness, accountability, and transparency. Bostrom and Yudkowsky (2014) claimed criteria of machine ethics related to the AI algorithm, such as responsibility or transparency. Arrieta et al. (2019) identified understandability, comprehensibility, and interpretability, for instance, whereas Fjeld et al. (2020) proposed eight core AI principles, including privacy, security, and so on. Whittlestone et al. (2019) also explored the use of ethical principles in research, with privacy as the word appearing most frequently. In their scoping review, Jobin et al. (2019) identified 11 ethical principles related to guidelines on ethical AI, such as (non-) maleficence. In addition, there are several initiatives concerned with the ethical implications of AI, such as AI4People, the Asilomar AI principles of the Future of Life Institute, the General Principles of IEEE (Institute of Electrical and Electronics Engineers) or the Ethical Principles published by the European Commission's European Group on Ethics in Science and New Technology (Floridi et al., 2018). Synthesizing those principles together in one cluster, we present the principles of AI ethics in Table N2. It should be noted that we chose our own classification based on the most common AI principles named by the authors.

Table N2
Principles of AI Ethics.

AI principles	Characteristics
Transparency & accountability (explainability, explicability, verifiability, understandability, interpretability, intelligibility, replicability, accuracy)	• Distribute accountability (Fjeld et al., 2020) • Provide information to understand outputs and scope of usage (Fjeld et al., 2020; Floridi et al., 2018) • Avoid non-transparent and unreliable outcomes (Leslie, 2019) • Demonstrate who accounts for the decision-making process and make decisions reviewable (Fjeld et al., 2020; Jobin et al., 2019) • Increase explainability and interpretability (Jobin et al., 2019)
Privacy and (non-) maleficence (security, safety, protection, no harm, precaution, prevention, integrity, incorruptibility, caution, reliability,)	• Avoid misuse, harm, and negative consequences (Floridi et al., 2018; Jobin et al., 2019) • Ensure that individuals' rights are respected (Fjeld et al., 2020; Leslie, 2019) • Avoid unwanted bias (Jobin et al., 2019; Leslie, 2019) • Ensure that AI performs as intended, being safe and resistant against unauthorized users (Fjeld et al., 2020)
Justice and fairness (consistency, inclusion, solidarity, equality, equity, diversity, plurality, accessibility, reversibility, remedy, redress, distribution, prosperity, non-bias, non-discrimination)	• Respect all interests by eliminating discrimination (Floridi et al., 2018) • Promote prosperity and solidarity (Floridi et al., 2018) • Maximize and promote inclusion/inclusiveness (Fjeld et al., 2020) • Provide fair access to AI (Jobin et al., 2019)
Beneficence & sustainability (well-being, dignity, peace, social or common good, benefits, long-term orientation)	• Promote well-being (Floridi et al., 2018) • Ensure positive benefits (Jobin et al., 2019) and long-term sustainable impact for mankind (Floridi et al., 2018) • Promote human values (Fjeld et al., 2020) • Ensure sustainable future developments (Jobin et al., 2019)
Responsibility & autonomy (freedom, consent, choice, self-determination, liberty, empowerment, liability, integrity)	• Ensure that important decisions remain by humans (Fjeld et al., 2020; Floridi et al., 2018) so that humans as the actor who is ultimately responsible (Jobin et al., 2019) • Ensure that individuals involved in the development and deployment of AI-based technologies are able to evaluate its impact and can be judged based on their professionalism (Fjeld et al., 2020) • Increase the balance between decision-making power of the *human agency* and the *material agency* (Floridi et al., 2018) • Ensure the protection of the intrinsic value of human choice and human self-determination (Floridi et al., 2018; Jobin et al., 2019; Leslie, 2019)

As our focus lies on the ethical managing human-AI interaction from both perspectives: AI interacting with humans (SQ1), but also humans interacting with AI (SQ2), we combine the principles of business ethics and AI ethics into categories of "human-AI ethics" in Table N3:

Table N3
Categories of Human-AI Ethics.

Categories Human-AI Ethics	Characteristics
Transparency & accountability (accuracy, explainability, explicability, distinctiveness, instrumentality, interpretability, intelligibility, reliability, replicability, understandability, visibility, verifiability)	*Duty ethics*: principles that ensure that processes of human-AI interaction are easy to understand, decisions are reviewable, and processes of human-AI interaction can be investigated dependent on the interests. *Virtue ethics:* the human's and AI's/developer's ideology that ensures that processes of human-AI interaction are easy to understand, decisions are reviewable, and processes of human-AI interaction can be investigated dependent on the interests.
Privacy and (non–) maleficence (caution, incorruptibility, integrity, no harm, precaution, prevention, protection, safety, security)	*Duty ethics*: principles that prevent personal privacy and security of humans during human-AI interaction and avoid misuse of AI through humans and vice versa. *Virtue ethics:* the human's and AI's/developer's ideology that aims to prevent personal privacy and security of humans during human-AI interaction and avoid misuse of AI through humans and vice versa.
Justice and fairness(accessibility, consensus, consistency, distribution, diversity, equality, equity, inclusion, non-bias, non-discrimination, plurality, prosperity, redress, relevance, remedy, reversibility, solidarity)	*Duty ethics:* principles that are based on the respect of all interests of humans to maximize inclusion and the equal treatment of humans in all types of human-AI interaction through adequate management processes. *Virtue ethics*: the human's and AI's/developer's ideology that aims to respect all interests of humans to maximize inclusion and the equal treatment of humans in all types of human-AI interaction through adequate management processes.
Beneficence & sustainability (benefits, dignity, long-term orientation, peace, social or common good, well-being)	*Duty ethics*: principles that aim for beneficial human-AI interactions with a long-term positive impact. *Virtue ethics*: the human's and AI's/developer's ideology that aims to achieve beneficial human-AI interactions with a long-term positive impact.
Responsibility & autonomy (agreement, authority, consent, choice, empowerment, freedom, legitimacy, liberty, liability, self-determination, validity)	*Duty ethics*: principles that balance the power of AI and humans in human-AI interaction or managing human-AI interaction as they are based on the agreement among all entities being involved in the human-AI interaction and its consequences. *Virtue ethics*: the human's and AI's/developer's ideology that aims to balance the power of AI and humans in human-AI interaction or to manage human-AI interaction as they are based on the agreement among all entities being involved in the human-AI interaction and its consequences.
Humanity (can be associated with the "uncanny valley" phenomenon in robotics following Park et al. (2021, p. 7) and Rajaobelina et al. (2021, p. 2351))	*Duty ethics*: principles that aim to consider possible disagreements between humans and AI linked to the transfer of ethical knowledge and experiences between humans and AI by integrating feedback loops or human interfaces for emotional support in the process of human-AI interaction. *Virtue ethics:* the human's and the manager's ideology that aims to use their experiences and emotional nature to integrate humanity, emotional support, and human values in human-AI interaction to reduce moral burden.

Appendix O. Strategic management

As mentioned above, strategic management contains the management of several strategies on three levels: the corporate, the business, and the functional level. The strategies behind those levels serve as a base for the keyword search (see Table O1).

Table O1
Levels of Strategic Management.

Levels	Description	Authors
Corporate Level	• Overarching corporate or global strategy • Strategic planning and coordination of several substrategies to gain competitive advantages (a rather external view) • Organizations goals, vision, and mission	(Henderson and Venkatraman, 1999; Karimi and Konsynski, 2002, pp. 92–93; Ritson, 2008, p. 13; Somogyi and Galliers, 2009, pp. 11–12)
Business Level	• Development of business strategies • Linked to the organizations' structure and internal processes (a rather internal view)	
Functional Level	• Development of strategies that are related to single departments of the organization	

We further identified four processes of strategic management given in Table O2:

Table O2
Processes of Strategic Management.

Processes	Description	Authors
Planning Processes	• Based on long-term planning horizons • Especially linked to the corporate strategy as it is about making formal strategic plans regarding the corporate strategy on the corporate level	Galliers and Sutherland, 2009, p. 39; Henderson and Venkatraman, 1999; Karimi and Konsynski, 2002, pp. 92–93; Peppard and Ward, 2004; Somogyi and Galliers, 2009, 8, 15
Strategic Processes	• Based on long-term planning horizons • Rather planning activities • Actions mainly on the business level to achieve the several business strategies, but also the corporate strategy (sometimes also on the corporate level)	
Tactic Processes	• Based on middle planning horizons • Coordination and control activities • Actions on the business level to evaluate business strategies and foster functional strategies by determining how best practices can be embedded in operational processes • Between the functional level and the business level	
Operational Processes	• Based on short planning horizons • Actions on the functional level to optimize operations by meeting the budget targets and being productive operatively • Daily working routines to achieve functional strategies	

References

Abdelaal, A., 2021. Grand Research Challenges Facing Ethically Aligned Artificial Intelligence. In: Proceedings of the 27th American Conference on Information Systems (AMCIS). Proceedings of the 27th American Conference on Information Systems (AMCIS 2021), Virtual. 09.08.2021-13.08.2021.

Acikgoz, Y., Davison, K.H., Compagnone, M., Laske, M., 2020. Justice perceptions of artificial intelligence in selection. Int. J. Select. Assess. (IJSA) 28, 399–416. https://doi.org/10.1111/ijsa.12306.

Alsheibani, S.A., Messom, D., Cheung, Y., Alhosni, M., 2020. Reimagining the Strategic Management of Artificial Intelligence: Five recommendations for business leaders, in: Proceedings of the 26th American Conference on Information Systems (AMCIS 2020). 26th Americas Conference on Information Systems (AMCIS 2020), Salt Lake City, Utah, USA. 10.08.2020-14.08.2020, pp. 1–10.

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., Horvitz, E., 2019. Guidelines for Human-AI Interaction. In: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI). CHI '19: CHI Conference on Human Factors in Computing Systems (CHI), Glasgow, Scotland. 04.05.2019.-09.05.2019, pp. 1–13.

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2019. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, 67 pp. https://arxiv.org/pdf/1910.10045v2.pdf (accessed 27 January 2020).

Asatiani, A., Malo, P., Nagbøl, P.R., Penttinen, E., Rinta-Kahila, T., Salovaara, A., 2020. Challenges of explaining the behavior of black-box AI systems. MIS Quart. Executive (MISQE) 259–278. https://doi.org/10.17705/2msqe.00037.

Association for Information Systems, 2011. Senior scholars' basket of journals. https://aisnet.org/page/SeniorScholarBasket (accessed 16 January 2022).

Australian Research Council, 2011. Excellence in research for Australia 2010 ranked conference list. https://webarchive.nla.gov.au/awa/20110217201708/http://www.arc.gov.au/pdf/ERA_report.pdf (accessed 27 January 2022).

Baidoo-Anu, D., Owusu Ansah, L., 2023. Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Available at SSRN 4337484.

Baird, A., Maruping, L.M., 2021. The next generation of research on IS use: A theoretical framework of delegation to and from agentics IS artifacts. MIS Quart. (MISQ) 45, 315–341. https://doi.org/10.25300/MISQ/2021/15882.

Bankins, S., Formosa, P., Griep, Y., Richards, D., 2022. AI Decision Making with Dignity? Contrasting Workers' Justice Perceptions of Human and AI Decision Making in a Human Resource Management Context. Inf. Syst. Front. 24 (3), 857–875. https://doi.org/10.1007/s10796-021-10223-8.

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M.T., Weld, D., 2021. Does the whole exceed its parts? The effect of AI explanations on complementary team performance, in: CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21: CHI Conference on Human Factors in Computing Systems (CHI 2021), Yokohama,Japan. 08.05.2021-13.05.2021. Association for Computing Machinery, New York, NY, USA, pp. 1–16.

Barad, K., 2003. Posthumanist Performativity: Toward an understanding of how matter comes to matter. Signs J. Women Cult. Soc. 28, 801–831. https://doi.org/10.1086/345321.

Baron, L.F., Gomez, R., 2016. The Associations between Technologies and Societies: The Utility of Actor-Network Theory. Science, Technol. Soc. 21, 129–148. https://doi.org/10.1177/0971721816640615.

Basu, S., Garimella, A., Han, W., Dennis, A., 2021. Human decision making in AI augmented systems: Evidence from the initial coin offering market, in: Proceedings of the 54th Hawaii International Conference on Systems (HICCS). 54th Hawaii International Conference on System Sciences (HICSS-54), Kauai, Hawaii. 05.01.2021, pp. 176–185.

Benbya, H., Davenport, T.H., Pachidi, S., 2020. Editors' comment: Artificial Intelligence in organizations: current state and future opportunities. MIS Quart. Execut. (MISQE) 19, 9–21. https://doi.org/10.2139/ssrn.3741983.

Berente, N., Gu, B., Recker, J., Santhanam, R., 2021. Managing Artificial Intelligence. MIS Quart. (MISQ) 45, 1433–1450. https://doi.org/10.25300/MISQ/2021/16274.

Boada, J.P., Maestre, B.R., Genís, C.T., 2021. The ethical issues of social assistive robotics: A critical literature review. Technol. Soc. 67 (101726), 1–13. https://doi.org/10.1016/j.techsoc.2021.101726.

Boell, S.K., Cecez-Kecmanovic, D., 2015. On being 'systematic' in literature reviews in IS. J. Inform. Technol. (JIT) 30, 161–173. https://doi.org/10.1057/jit.2014.26.

Bonn, I., Fisher, J., 2005. Corporate Governance and Business Ethics: Insights from the strategic planning experience. Corp. Gov. 13, 730–738. https://doi.org/10.1111/j.1467-8683.2005.00466.x.

Bostrom, N., Yudkowsky, E., 2014. The ethics of Artificial Intelligence. In: Frankish, K., Ramsey, W.M. (Eds.), The Cambridge Handbook of Artificial Intelligence. Cambridge University Press, pp. 316–334.

Bowie, N.E., 1998. A kantian theory of meaningful work. J. Bus. Ethics (JBE) 17, 1083–1092. https://doi.org/10.1023/A:1006023500585.
Burton-Jones, A., Butler, B.S., Scott, S., Xu, S.X., 2021. Next-generation information systems theorizing: A call to action. MIS Quart. (MISQ) 45, 301–314. https://doi.org/10.25300/MISQ/2021/15434.
Carter, L., Liu, D., Cantrell, C., 2020. Exploring the intersection of the digital divide and Artificial Intelligence: A hermeneutic Literature Review. AIS Trans. Human-Comput. Interaction (THCI) 12, 253–275. https://doi.org/10.17705/1thci.00138.
Cecez-Kecmanovic, D., Galliers, R.D., Henfridsson, O., Newell, S., Vidgen, R., 2014. The Sociomaterialty of Information Systems: Current Status, Future Directions. MIS Quart. (MISQ) 38 (3), 809–830. https://doi.org/10.25300/MISQ/2014/38:3.3.
Chao, C.-H., 2019. Ethics issues in Artificial Intelligence, in: 24th International Conference on Technologies and Applications of Artificial Intelligence (TAAI 2019). 24th International Conference on Technologies and Applications of Artificial Intelligence (TAAI 2019), Kaohsiung, Taiwan. 21.11.2019-23.11.2019, pp. 1–6.
Churchill, L.R., 1999. Are we professionals? A critical look at the social role of bioethicists. Daedalus 128, 253–274.
Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. Educ. Psychol. Meas. 20 (1), 37–46.
Coldwell, D.A., Billsberry, J., van Meurs, N., Marsh, P., 2008. The effects of person-organization ethical fit on employee attraction and retention: Towards a testable explanatory model. J. Bus. Ethics (JBE) 78, 611–622. https://doi.org/10.1007/s10551-007-9371-y.
Contractor, N., Monge, P., Leonardi, P.M., 2011. Multidimensional Networks and the Dynamics of Sociomateriality: Bringing Technology Inside the Network. Int. J. Commun. 5, 682–720.
Coombs, C., Stacey, P., Kawalek, P., Simeonova, B., Becker, J., Bergener, K., Carvalho, J., Fantinato, M., Garmann-Johnsen, N.F., Grimme, C., Stein, A., Trautmann, H., 2021. What is it about humanity that we can't give away to intelligent machines? A European perspective. Int. J. Inform. Manage. (IJIM) 58 (102311), 1–13. https://doi.org/10.1016/j.ijinfomgt.2021.102311.
Cooren, F., 2004. Textual Agency: How texts do things in organizational settings. Organization 11, 373–393. https://doi.org/10.1177/1350508404041998.
Cranefield, J., Doyle, C., 2022. "You have a few wordy spots": Role-related impacts of working with an intelligent assistant, in: ECIS 2022 Research Paper. 30th European Conference on Information Systems (ECIS), Timisoara. 18.06.2022-24.06.2022, pp. 1–16.
Daily, M., Oulasvirta, A., Rekimoto, J., 2017. Technology for Human Augmentation. Computer 50, 12–15. https://doi.org/10.1109/MC.2017.39.
Dando, N., Swift, T., 2003. Transparency and Assurance: Minding the credibility gap. J. Bus. Ethics (JBE) 44, 195–200. https://doi.org/10.1023/A:1023351816790.
Davenport, T.H., Kirby, J., 2016. Just how smart are smart machines? MITSloan Manage. Rev. 57, 21–25.
Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., Ebel, P., 2019a. The future of Human-AI collaboration: A taxonomy of design knowledge for hybrid Intelligence Systems, in: 52nd Hawaii International Conference on System Sciences (HICSS). 52nd Hawaii International Conference on System Sciences (HICSS), Maui, Hawaii. 08.01.2019-11.01.2019, pp. 1–10.
Dellermann, D., Ebel, P., Söllner, M., Leimeister, J.M., 2019b. Hybrid Intelligence. Bus. Inform. Syst. Eng. (BISE) 61, 637–643. https://doi.org/10.1007/s12599-019-00595-2.
Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdullah, A.M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., Chowdhury, S., Crick, T., Cunningham, S.W., Davies, G.H., Davison, R.M., Dé, R., Dennehy, D., Duan, Y., Dubey, R., Dwivedi, R., Edwards, J.S., Flavián, C., Gauld, R., Grover, V., Hu, M.-C., Janssen, M., Jones, P., Junglas, I., Khorana, S., Kraus, S., Larsen, K.R., Latreille, P., Laumer, S., Malik, F.T., Mardani, A., Mariani, M., Mithas, S., Mogaji, E., Nord, J.H., O'Connor, S., Okumus, F., Pagani, M., Pandey, N., Papagiannidis, S., Pappas, I.O., Pathak, N., Pries-Heje, J., Raman, R., Rana, N.P., Rehm, S.-V., Ribeiro-Navarrete, S., Richter, A., Rowe, F., Sarker, S., Stahl, B.C., Tiwari, M.K., van der Aalst, W., Venkatesh, V., Viglia, G., Wade, M., Walton, P., Wirtz, J., Wright, R., 2023. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int. J. Inf. Manag. 71, 102642 https://doi.org/10.1016/j.ijinfomgt.2023.102642.
Ehsan, U., Riedl, M.O., 2020. Human-centered explainable AI: Towards a reflective sociotechnical approach, in: 22nd International Conference on Human-Computer Interaction (HCI). 22nd International Conference on Human-Computer Interaction (HCI), Kopenhagen, Dänemark. 19.07.2020-24.07.2020, pp. 1–19.
Eitel-Porter, R., 2021. Beyond the Promise: Implementing Ethical AI. AI Ethics 1, 73–80. https://doi.org/10.1007/s43681-020-00011-6.
Etzioni, A., Etzioni, O., 2017. Incorporating Ethics into Artificial Intelligence. J. Ethics 21, 403–418. https://doi.org/10.1007/s10892-017-9252-2.
Faraj, S., Pachidi, S., Sayegh, K., 2018. Working and organizing in the age of the learning algorithm. Inf. Organ. 28, 62–70. https://doi.org/10.1016/j.infoandorg.2018.02.005.
Faulkner, P., Runde, J., 2011. The social, the material, and the ontology of non-material technological objects. European Group for Organizational Studies (EGOS) -, 1–34.
Figueroa-Armijos, M., Clark, B.B., Da Motta Veiga, S.P., 2022. Ethical perceptions of AI in hiring and organizational trust: The role of performance expectancy and Social Influence. J. Bus. Ethics (JBE) 1–19. https://doi.org/10.1007/s10551-022-05166-2.
Financial Times, 2016. 50 journals used in FT research rank. https://www.ft.com/content/3405a512-5cbb-11e1-8f1f-00144feabdc0 (accessed 16 January 2022).
Fjeld, J., Achten, N., Hilligoss, H.: Nagy, A., Srikumar, M., 2020. Principled Artificial Intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI, 39 pp. http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420 (accessed 27 January 2022).
Flathmann, C., Schelble, B.G., Zhang, R., McNeese, N.J., 2021. Modeling and guiding the creation of ethical Human-AI teams, in: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society. AIES '21: AAAI/ACM Conference on AI, Ethics, and Society, Virtual Event USA. 19.05.2021-21.05.2021. ACM, New York, NY, USA, pp. 469–479.
Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., Vayena, E., 2018. AI4People-An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Mind. Mach. 28, 689–707. https://doi.org/10.1007/s11023-018-9482-5.
Fuegener, A., Grahl, J., Gupta, A., Ketter, W., 2021. Cognitive challenges in Human-AI collaboration: investigating the path towards productive delegation. Inform. Syst. Res. (ISR) 33, 678–696. https://doi.org/10.1287/isre.2021.1079.
Galliers, R.D., Sutherland, A.R., 2009. The evolving Information Systems strategy. In: Galliers, R., Leidner, D.E. (Eds.), Challenges and strategies in managing Information Systems, 4th ed. Routledge, New York, pp. 1–31.
Gandz, J., Hayes, N., 1988. Teaching Business Ethics. J. Bus. Ethics (JBE) 7 (9), 657–669.
Gasser, U., Almeida, V., 2017. A layered model for AI governance. IEEE Internet Comput. 21, 58–62. https://doi.org/10.1109/MIC.2017.4180835.
Giermindl, L.M., Strich, F., Christ, O., Leicht-Deobald, U., Redzepi, A., 2022. The dark sides of people analytics: Reviewing the perils for organisations and employees. Eur. J. Inform. Syst. (EJIS) 31, 410–435. https://doi.org/10.1080/0960085X.2021.1927213.
Goodpaster, K.E., 1991. Business Ethics and stakeholder analysis. Bus. Ethics Quart. (BEQ) 1, 53–73. https://doi.org/10.2307/3857592.
Greene, T., Shmueli, G., Ray, S., forthcoming. Taking the Person Seriously: Ethically-aware IS Research in the Era of Reinforcement Learning-based Personalization. JAIS Preprints (Journal of the Association for Information Systems), 1–35. https://doi.org/10.17705/1jais.00491.
Greenwood, M.R., 2002. Ethics and HRM: A review and conceptual analysis. J. Bus. Ethics (JBE) 36, 216–278. https://doi.org/10.1023/A:1014090411946.
Grønsund, T., Aanestad, M., 2020. Augmenting the algorithm: Emerging human-in-the-loop work configurations. J. Strateg. Inf. Syst. 29 (101614), 1–16. https://doi.org/10.1016/j.jsis.2020.101614.
Gupta, S., Kamboj, S., Bag, S., 2021. Role of risks in the development of responsible Artificial Intelligence in the digital healthcare domain. Information Systems Frontiers (ISF) -, 1–18.
Hamilton, C., Swart, W., Stokes, G.M., 2021. Developing a measure of social, ethical, and legal content for intelligent cognitive assistants. J. Strateg. Innov. Sustain. 16, 1–37. https://doi.org/10.33423/jsis.v16i3.
Han, S., Kelly, E., Nikou, S., Svee, E.O., 2020. Reflections on Artificial Intelligence alignment with human values: A phenomenological perspective, in: 28th European Conference on Information Systems 2020 (ECIS). 28th European Conference on Information Sytems (ECIS), Marrakech, Morocco. 15.06.2020-17.06.2020, pp. 1–17.
Harzing, A.-W., 2020. Journal quality list, 24 pp. https://harzing.com/.
Hasnas, J., 2013. Teaching Business Ethics: The principles approach. J. Bus. Ethics Education (JBEE) 10, 275–304. https://doi.org/10.5840/jbee20131014.

Henderson, J.C., Venkatraman, H., 1999. Strategic alignment: Leveraging information technology for transforming organizations. IBM Syst. J. 38, 472–484. https://doi.org/10.1147/SJ.1999.5387096.

Heyder, T., Posegga, O., 2021. Extending the foundations of AI literacy, in: Proceedings of the 42nd International Conference on Information Systems (ICIS), Austin, Texas.

Hinsen, S., Hofmann, P., Jöhnk, J., Urbach, N., 2022. How can organizations design purposeful Human-AI interactions: A practical perspective from existing use cases and interviews, in: Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS). 55th Hawaii International Conference on System Sciences (HICSS-55), Maui, Hawaii (Virtual). 04.01.2022-07.01.2022. University of Hawaii at Manoa Hamilton Library, Honolulu, HI, pp. 1–11.

Hoeddinghaus, M., Sondern, D., Hertel, G., 2021. The automation of leadership functions: Would people trust decision algorithms? Comput. Hum. Behav. 116, 1–57. https://doi.org/10.1016/j.chb.2020.106635.

Hohenstein, J., Jung, M., 2020. AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. Comput. Hum. Behav. 106, 1–13. https://doi.org/10.1016/j.chb.2019.106190.

Holstein, K., Wortman Vaughan, J., Daumé, H., Dudik, M., Wallach, H., 2019. Improving fairness in machine learning systems, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19: CHI Conference on Human Factors in Computing Systems (CHI), Glasgow, Scotland. 04.05.2019-09.05.2019. ACM, New York, NY, USA, pp. 1–16.

Hosmer, L.T., 1987. Ethical Analysis and Human Resource Management. Hum. Resour. Manage. 26, 313–330. https://doi.org/10.1002/hrm.3930260302.

Islam, G., Greenwood, M., 2021. Reconnecting to the social in Business Ethics. J. Bus. Ethics (JBE) 170, 1–4. https://doi.org/10.1007/s10551-021-04775-7.

Janssen, C.P., Donker, S.F., Brumby, D.P., Kun, A.L., 2019. History and future of human-automation interaction. Int. J. Hum Comput Stud. 131, 99–107. https://doi.org/10.1016/j.ijhcs.2019.05.006.

Jia, H., Wu, M., Sundar, S.S., 2022. Do we blame it on the machine? Task outcome and agency attribution in human-technology collaboration, in: Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS). 55th Hawaii International Conference on System Sciences (HICSS-55), Maui, Hawaii (Virtual). 04.01.2022-07.01.2022. University of Hawaii at Manoa Hamilton Library, Honolulu, HI, pp. 388–397.

Jiang, J., Adam, M., Benlian, A., 2021. Algoactivistic practices in ridesharing - A topic modeling & grounded theory approach, in: Proceedings of the 29th European Conference on Information Systems (ECIS). 29th European Conference on Information Systems (ECIS), Marrakech, Morocco. 14.06.2021-16.06.2021, pp. 1–17.

Jobin, A., Ienca, M., Vayena, E., 2019. Artificial Intelligence: The global landscape of AI ethics guidelines. Nat. Mach. Intell. 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2.

Jones, C., Parker, M., ten Bos, R., 2005. For Business Ethics. Routledge, London, UK, p. 223.

Jones, M., 2014. A matter of life and death: exploring conceptualizations of sociomateriality in the context of critical care. MIS Quart. (MISQ) 38, 895–926. https://doi.org/10.25300/MISQ/2014/38.3.12.

Kamar, E., 2016. Directions in hybrid intelligence: Complementing AI Systems with human intelligence, in: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI). Twenty-Fifth International Joint Conferences on Artificial Intelligence (IJCAI), New York, NY. 09.07.2016-15.07.2016, pp. 4070–4073.

Kant, I., 1991. The Metaphysics of Morals. Cambridge University Press, Cambridge.

Karimi, J., Konsynski, B.R., 2002. The Information Technology and management infrastructure strategy. In: Galliers, R.D., Leidner, D.E. (Eds.), Strategic Information Management: Challenges and strategies in managing Information Systems, 3rd Edition, 3rd ed. Routledge, London, UK, pp. 89–112.

Kelley, S., 2022. Employee perceptions of the effective adoption of AI principles. J. Bus. Ethics (JBE) 178, 871–893. https://doi.org/10.1007/s10551-022-05051-y.

Kitchenham, B., Charters, S., 2007. Guidelines for performing systematic literature reviews in software engineering: Version 2.3, 65 pp. https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf.

Klumpp, M., 2018. Automation and Artificial Intelligence in business logistics systems: human reactions and collaboration requirements. Int. J. Log. Res. Appl. 21, 224–242. https://doi.org/10.1080/13675567.2017.1384451.

Klumpp, M., Zijm, H., 2019. Logistics innovation and social sustainability: How to prevent an artificial divide in Human-Computer Interaction. J. Bus. Logist. (JBL) 40, 265–278. https://doi.org/10.1111/jbl.12198.

Kolbjørnsrud, V., Amico, R., Thomas, R.J., 2017. Partnering with AI: how organizations can win over skeptical managers. Strategy Lead (SL) 45, 37–43. https://doi.org/10.1108/SL-12-2016-0085.

Larsen, K.R., Hovorka, D.S., Dennis, A.R., West, J.D., 2019. Understanding the Elephant: The Discourse Approach to Boundary Identification and Corpus Construction for Theory Review Articles. J. Assoc. Inform. Syst. (JAIS) 20, 887–927. https://doi.org/10.17705/1jais.00556.

Latour, B., 1992. Where are the missing masses? The sociology of a few mundane artifacts. Shaping technology/building society: Studies in sociotechnical change 1, 225–258.

Latour, B., 2005. Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford University Press, Oxford, New York, p. 301.

Law, J., 1992. Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. Systems Practice 5, 379–393. https://doi.org/10.1007/BF01059830.

Lee, M.K., 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. Big Data & Society (BD&S) 5, 1–16. https://doi.org/10.1177/2053951718756684.

Lee, M.K., Kusbit, D., Metsky, E., Dabbish, L., 2015. Working with machines, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15: CHI Conference on Human Factors in Computing Systems (CHI), Seoul Republic of Korea. 18.04.2015-23.04.2015. ACM, New York, NY, USA, pp. 1603–1612.

Leonardi, P.M., 2011. When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. MIS Quarterly (MISQ) 35, 147–167. https://doi.org/10.2307/23043493.

Leonardi, P.M., 2012. Materiality, Sociomateriality, and Socio-Technical Systems: What do these terms mean? How are they related? Do we need them? In: Leonardi, P.M., Nardi, B.A., Kallinikos, J. (Eds.), Materiality and organizing: Social Interaction in a technological world. Oxford University Press, Oxford, UK, pp. 25–48.

Leslie, D., 2019. Understanding Artificial Intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. The Alan Turing Institute 97.

Liao, Q.V., Gruen, D., Miller, S., 2020. Questioning the AI: Informing Design Practices for Explainable AI User Experiences, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20: CHI Conference on Human Factors in Computing Systems (CHI), Honolulu HI USA. 25.04.2020-30.04.2020. ACM, New York, NY, USA, pp. 1–15.

Long, D., Magerko, B., 2020. What is AI literacy? Competencies and design considerations, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20: CHI Conference on Human Factors in Computing Systems (CHI), Honolulu HI USA. 25.04.2020-30.04.2020. ACM, New York, NY, USA, pp. 1–16.

Lowry, P.B., Moody, G.D., Gaskin, J., Galletta, D.F., Humpherys, S.L., Barlow, J.B., Wilson, D.W., 2013. Evaluating journal quality and the association for information systems senior scholars' journal basket via bibliometric measures: Do expert journal assessments add value? MIS Quarterly (MISQ) 37 (4), 993–1012. https://doi.org/10.25300/MISQ/2013/37.4.01.

MacInnis, D.J., 2011. A framework for conceptual contributions in marketing. J. Market. (JM) 75, 136–154. https://doi.org/10.1509/jmkg.75.4.136.

MacKenzie, S.B., 2003. The Dangers of Poor Construct Conceptualization. J. Acad. Mark. Sci. 31, 323–326. https://doi.org/10.1177/0092070303031003011.

Madaio, M.A., Stark, L., Wortman Vaughan, J., Wallach, H., 2020. Co-Designing checklists to understand organizational challenges and opportunities around fairness in AI, in: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI). CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI), Honolulu, USA. 25.04.2020-30.04.2020. Association for Computing Machinery, New York, NY, USA, pp. 1–14.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., Söllner, M., 2019. AI-Based digital assistants. Bus. Inform. Syst. Eng. (BISE) 61, 535–544. https://doi.org/10.1007/s12599-019-00600-8.

Malik, N., Tripathi, S.N., Kar, A.K., Gupta, S., 2021. Impact of artificial intelligence on employees working in industry 4.0 led organizations. Int. J. Manpower (IJM) 43, 334–354. https://doi.org/10.1108/IJM-03-2021-0173.

Marabelli, M., Newell, S., Handunge, V., 2021. The lifecycle of algorithmic decision-making systems: Organizational choices and ethical challenges. J. Strateg. Inf. Syst. 30 (3), 101683. https://doi.org/10.1016/j.jsis.2021.101683.
Markus, M.L., Rowe, F., 2021. Guest editorial: Theories of digital transformation: A progress report. J. Assoc. Inform. Syst. (JAIS) 22, 273–280. https://doi.org/10.17705/1jais.00661.
Markus, M.L., Rowe, F., 2023. The Digital Transformation Conundrum: Labels, Definitions, Phenomena, and Theories. JAIS 24, 328–335. https://doi.org/10.17705/1jais.00809.
Mayer, A.S., Haimerl, A., Strich, F., Fiedler, M., 2021. How corporations encourage the implementation of AI ethics, in: ECIS 2021 Research Papers. 29th European Conference on Information Systems (ECIS), Marrakech, Morocco. 14.06.2021-16.06.2021, 1-15.
Mayer, A.-S., Strich, F., Fiedler, M., 2020. Unintended consequences of introducing AI systems for decision making. MIS Quart. Executive (MISQE) 19, 239–257. https://doi.org/10.17705/2msqe.00036.
Mayring, P., 2014. Qualitative content analysis: Theoretical foundation, basic procedures and software solution. Klagenfurt.
McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochemia medica 22, 276–282. https://doi.org/10.1201/b16720.
Mingers, J., Harzing, A.W., 2007. Ranking journals in business and management: A statistical analysis of the Harzing data set. Eur. J. Inform. Syst. (EJIS) 16, 303–316. https://doi.org/10.1057/palgrave.ejis.3000696.
Mirbabaie, M., Brünker, F., Möllmann Frick, N., Stieglitz, S., 2022. The rise of artificial intelligence – understanding the AI identity threat at the workplace. Electron. Mark. 32, 73–99. https://doi.org/10.1007/s12525-021-00496-x.
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D., 2015. Human-level control through deep reinforcement learning. Nature 518, 529–533. https://doi.org/10.1038/nature14236.
Mollick, E., 2022. ChatGPT is a tipping point for AI. Harv. Bus. Rev.
Mutch, A., 2013. Sociomateriality — Taking the wrong turning? Inf. Organ. 23, 28–40. https://doi.org/10.1016/j.infoandorg.2013.02.001.
Ng, D., Leung, J., Chu, S., Qiao, M.S., 2021. Conceptualizing AI literacy: An exploratory review. Comput. Educ.: Artif. Intell. 2, 1–11. https://doi.org/10.1016/j.caeai.2021.100041.
Norman, D., 2017. Design, business models, and human-technology teamwork: Everything as a service. Res.-Technol. Management (RTM) 60, 26–30. https://doi.org/10.1080/08956308.2017.1255051.
Orlikowski, W.J., 2007. Sociomaterial practices: Exploring technology at work. Organ. Stud. (OS) 28, 1435–1448. https://doi.org/10.1177/0170840607081138.
Orlikowski, W.J., 2010. The sociomateriality of organisational life: considering technology in management research. Camb. J. Econ. (CJE) 34, 125–141. https://doi.org/10.1093/cje/bep058.
Orlikowski, W.J., Scott, S.V., 2008. 10 Sociomateriality: Challenging the separation of technology, work and organization. Acad. Manage. Ann. (Annals) 2, 433–474. https://doi.org/10.1080/19416520802211644.
O'Sullivan, S., Nevejans, N., Allen, C., Blyth, A., Leonard, S., Pagallo, U., Holzinger, K., Holzinger, A., Sajid, M.I., Ashrafian, H., 2019. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. Int. J. Med. Rob. Comput. Assisted Surg. 15 (1), e1968. https://doi.org/10.1002/rcs.1968.
Paré, G., Trudel, M.-C., Jaana, M., Kitsiou, S., 2015. Synthesizing information systems knowledge: A typology of literature reviews. Inf. Manag. 52, 183–199. https://doi.org/10.1016/j.im.2014.08.008.
Park, H., Ahn, D., Hosanagar, K., Lee, J., 2021. Human-AI Interaction in Human Resource Management: Understanding why employees resist algorithmic evaluation at workplaces and how to mitigate burdens, in: CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21: CHI Conference on Human Factors in Computing Systems (CHI 2021), Yokohama, Japan. 08.05.2021-13.05.2021. Association for Computing Machinery, New York, New York, pp. 1–15.
Parker, M., 2002. Against Management: Organization in the age of managerialism. Oxford University Press, Oxford, UK.
Pavlik, J.V., 2023. Collaborating With ChatGPT: Considering the Implications of Generative Artificial Intelligence for Journalism and Media Education. J. Mass Commun. Educator 78, 84–93. https://doi.org/10.1177/10776958221149577.
Payne, S.L., Wayland, R.F., 1999. Ethical obligation and diverse values assumptions in HRM. Int. J. Manpower (IJM) 20, 297–308. https://doi.org/10.1108/01437729910286083.
Peppard, J., Ward, J., 2004. Beyond strategic information systems: towards an IS capability. J. Strateg. Inf. Syst. 13, 167–194. https://doi.org/10.1016/j.jsis.2004.02.002.
Polyviou, A., Zamani, E.D., 2022. Are we Nearly There Yet? A Desires & Realities Framework for Europe's AI Strategy. Inf. Syst. Front. 1–17 https://doi.org/10.1007/s10796-022-10285-2.
Pregenzer, M., Remus, U., Wiener, M., 2021. Algorithms in the driver's seat: explaining workers' reactions to algorithmic control, in: Proceedings of the 29th European Conference on Information Systems (ECIS). 29th European Conference on Information Systems (ECIS), Marrakech, Morocco. 14.06.2021-16.06.2021, pp. 1–15.
Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., 2018. Improving language understanding by generative pre-training. Preprint.
Rai, A., Constantinides, P., Sarker, S., 2019. Next generation digital platforms: Toward Human-AI hybrids. MIS Quarterly (MISQ) 43, 3–9.
Raisamo, R., Rakkolainen, I., Majaranta, P., Salminen, K., Rantala, J., Farooq, A., 2019. Human augmentation: Past, present and future. Int. J. Hum Comput Stud. 131, 131–143. https://doi.org/10.1016/j.ijhcs.2019.05.008.
Raisch, S., Krakowski, S., 2020. Artificial Intelligence and Management: The automation-augmentation paradox. Acad. Manag. Rev. 46, 192–210. https://doi.org/10.5465/2018.0072.
Rajaobelina, L., Prom Tep, S., Arcand, M., Ricard, L., 2021. Creepiness: Its antecedents and impact on loyalty when interacting with a chatbot. Psychol. Mark. 38, 2339–2356. https://doi.org/10.1002/mar.21548.
Rakowski, R., Polak, P., Kowalikova, P., 2021. Ethical Aspects of the Impact of AI: the Status of Humans in the Era of Artificial Intelligence. Soc 58 (3), 196–203. https://doi.org/10.1007/s12115-021-00586-8.
Reckwitz, A., 2002. Toward a theory of social practices: A development in culturalist theorizing. Eur. J. Social Theory (EJST) 5, 243–263. https://doi.org/10.1177/13684310222225432.
Rezaee, Z., 2008. Corporate governance and ethics. John Wiley & Sons, Boston, MA.
Rieder, T.N., Hutler, B., Mathews, D., 2020. Artificial Intelligence in service of human needs: Pragmatic first steps toward an ethics for semi-autonomous agents. AJOB Neurosci. 11, 120–127. https://doi.org/10.1080/21507740.2020.1740354.
Ritson, N., 2008. Strategic Management, 2nd ed. bookboon.com, p. 117 pp..
Rivard, S., 2021. Theory building is neither an art nor a science: It is a craft. J. Inform. Technol. (JIT) 36, 316–328. https://doi.org/10.1177/0268396220911938.
Rix, J., Hess, T., 2022. Hello, Mate! Insights from the field on leveraging machine teammates in organizations, in: Proceedings of the 26th Pacific Asia Conference on Information Systems (PACIS). 26th Pacific Asia Conference on Information Systems (PACIS), Virtual Event / Taipei, Taiwan / Sydney, Australia. 05.07.2022-09.07.2022.
Sauppé, A., Mutlu, B., 2015. The social impact of a robot co-worker in industrial settings, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15: 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI), Seoul Republic of Korea. 18.04.2015-23.04.2015. ACM, New York, NY, USA, pp. 3613–3622.
Scherer, M.U., 2016. Regulating Artificial Intelligence systems: Risks, challenges, competencies, and strategies. Harvard J. Law Technol. 29, 354–398. https://doi.org/10.2139/ssrn.2609777.
Schwartz, M.S., Saiia, D., 2012. Should firms go beyond profits? Milton Friedman versus broad CSR. Bus. Soc. Rev. 117, 1–13. https://doi.org/10.1111/j.1467-8594.2011.00397.x.
Schwarz, A., Mehta, M., Johnson, N., Chin, W.W., 2007. Understanding Frameworks and Reviews: A Commentary to Assist us in Moving Our Field Forward by Analyzing Our Past. ACM SIGMIS Database: the DATABASE for Advances in Information Systems 38 (3), 29–50.

Shahriari, B., Swersky, K., Wang, Z., Adams, R.P., de Freitas, N., 2016. Taking the human out of the loop: A review of bayesian optimization. Proc. IEEE 104, 148–175. https://doi.org/10.1109/JPROC.2015.2494218.

Shin, D., Park, Y.J., 2019. Role of fairness, accountability, and transparency in algorithmic affordance. Comput. Hum. Behav. 98, 277–284. https://doi.org/10.1016/j.chb.2019.04.019.

Shneiderman, B., 2020. Human-centered Artificial Intelligence: Three fresh ideas. AIS Trans. Human-Comput. Int. 12, 109–124. https://doi.org/10.17705/1thci.00131.

Siau, K., Wang, W., 2020. Artificial Intelligence (AI) ethics: Ethics of AI and ethical AI. Journal of Database Management (JDM) 31, 74–87. https://doi.org/10.4018/JDM.2020040105.

Someh, I., Wixom, B.H., Beath, C.M., Zutavern, A., 2022. Building an Artificial Intelligence explanation capability. MIS Quart. Executive (MISQE) 21 (5), 143–163. https://doi.org/10.17705/2msqe.00063.

Somogyi, E.K., Galliers, R.D., 2009. Developments in the application of Information Technology in business. In: Galliers, R., Leidner, D.E. (Eds.), Challenges and strategies in managing Information Systems, 4th ed. Routledge, New York, pp. 3–22.

Stevens, B., 1999. Communicating ethical values: A study of employee perceptions. Journal of Business Ethics (JBE) 20, 113–120.

Strich, F., Mayer, A.-S., Fiedler, M., 2021. What do I do in a world of Artificial Intelligence? Investigating the impact of substitutive decision-making AI systems on employees' professional role identity. J. Assoc. Inform. Syst. (JAIS) 22, 304–324. https://doi.org/10.17705/1jais.00663.

Te'eni, D., Avital, M., Hevner, A., Schoop, M., Schwartz, D.G., 2019. It takes two to tango: Choreographing the interactions between Human and Artificial Intelligence, in: 27th European Conference on Information Systems: Information Systems for a Sharing Society (ECIS). 27th European Conference on Information Systems: Information Systems for a Sharing Society (ECIS), Stockholm, Sweden. 08.06.2019-14.06.2019.

Templier, M., Paré, G., 2018. Transparency in literature reviews: An assessment of reporting practices across review types and genres in top IS journals. Eur. J. Inform. Syst. (EJIS) 27, 503–550. https://doi.org/10.1080/0960085X.2017.1398880.

Thompson, M., 2012. People, practice, and technology: Restoring Giddens' broader philosophy to the study of information systems. Inf. Organ. 22 (3), 188–207. https://doi.org/10.1016/j.infoandorg.2012.04.001.

VHB, 2011. JOURQUAL-3 Gesamtranking. https://vhbonline.org/vhb4you/vhb-jourqual/vhb-jourqual-3/tabellen-zum-download (accessed 27 January 2022).

Vial, G., 2019. Understanding digital transformation: A review and a research agenda. J. Strateg. Inf. Syst. 28, 118–144. https://doi.org/10.1016/j.jsis.2019.01.003.

Wang, Y., Xiong, M., Olya, H., 2020. Toward an understanding of responsible Artificial Intelligence practices, in: Proceedings of the 53rd Hawaii International Conference on System Sciences (HICSS). 53rd Hawaii International Conference on System Sciences 2020 (HICSS-53), Maui, Hawaii. 07.01.2020-10.01.2020, pp. 4962–4971.

Watson, G.J., Desouza, K.C., Ribiere, V.M., Lindič, J., 2021. Will AI ever sit at the C-suite table? The future of senior leadership. Bus. Horiz. 64, 465–474. https://doi.org/10.1016/j.bushor.2021.02.011.

Weber, R., 2012. Evaluating and Developing Theories in the Information Systems Discipline. J. Assoc. Inform. Syst. (JAIS) 13, 1–30. https://doi.org/10.17705/1jais.00284.

Whittlestone, J., Nyrup, R., Alexandrova, A., 2019. Ethical and societal implications of algorithms, data, and artificial intelligence: A roadmap for research. https://www.nuffieldfoundation.org/sites/default/files/files/Ethical-and-Societal-Implications-of-Data-and-AI-report-Nuffield-Foundat.pdf (accessed 27 January 2022).

Wilson, J., Daugherty, P.R., 2018. Collaborative Intelligence: Humans and AI are joining forces. Harv. Bus. Rev. 96, 114–123.

Winstanley, D., Woodall, J., Heery, E., 1996. Business ethics and human resource management: Themes and issues. Pers. Rev. 25, 5–12. https://doi.org/10.1108/00483489610148491.

Zhang, X., Zhao, Y., Tang, X., Zhu, H., Xiong, H., 2020. Developing fairness rules for talent intelligence management system, in: Proceedings of the 53rd Hawaii International Conference on System Sciences (HICSS). 53rd Hawaii International Conference on System Sciences (HICSS-53), Maui, Hawaii. 07.01.2020-10.01.2020, pp. 5882–5891.