2024

Analyzing The Use of Ethical Theories Within AI Ethics Research: A Systematic Scoping Review

Alina Hafner
Technical University of Munich, Germany, alina.hafner@tum.de

Teresa Hammerschmidt
University of Bamberg, Germany, teresa.heyder@uni-bamberg.de

Katharina Stolz
University of Stuttgart, Germany, katharina.stolz@bwi.uni-stuttgart.de

Analyzing The Use of Ethical Theories Within AI Ethics Research: A Systematic Scoping Review

Research Paper

Alina Hafner[1], Teresa Hammerschmidt[2], Katharina Stolz[3]

[1] Technical University of Munich, TUM School of Management, Heilbronn, Germany
alina.hafner@tum.de
[2] University of Bamberg, Information Systems and Social Networks, Bamberg, Germany
teresa.heyder@uni-bamberg.de
3 University of Stuttgart, Innovation and Service Management, Stuttgart, Germany
katharina.stolz@bwi-uni-stuttgart.de

Abstract. Artificial Intelligence (AI) ethics research is a multifaceted field, requiring different theoretical justifications in which researchers can ground their underlying perspectives on ethics. We provide an overview of the major normative ethical theories used in Information Systems research on AI ethics. Through a systematic scoping review, we assess the prevailing theories, their progress, and areas needing further study. Our findings reveal a dominance of deontological ethics, which results in determining ethics mainly from the AI's perspective by discussing ethical design principles but not from how a human user's virtue ethics perspective guides humans' moral behavior when collaborating with AI equally. We suggest that researchers recognize how normative ethical theories might bind their work, impacting their understanding of moral agency and responsibility and guiding Corporate Digital Responsibility practices for organizations striving for responsible AI design, deployment, and usage.

Keywords: Scoping review, AI ethics, ethical theories, digital responsibility

1 Introduction

With Artificial Intelligence (AI) able to the ability "to perform cognitive functions that we associate with human minds"(Rai et al., 2019, p. iii), a novel generation of technologies arises, exhibiting greater autonomy and complexity compared to previous ones (Baird & Maruping, 2021; Brendel et al., 2021). When AI informs, controls, or even automates specific tasks and decisions, questions of morality and responsibility arise – especially when unintended outcomes, such as the discrimination of certain groups of people due to data biases, occur (Giermindl et al., 2022; Kellogg et al., 2020). Novel forms of human-AI collaboration present several ethical challenges for individuals, organizations, and society (Dwivedi et al., 2023). This leads to a rising academic discourse in the context of Corporate Digital Responsibility (CDR) concerning responsibility questions that arise with the widespread use of autonomous technologies (Mihale-

Wilson et al., 2022) to form a transformation that is "not only … driven by what is technologically possible, but also what is socially desirable and sustainable." (Mueller, 2022, p. 689). CDR and AI ethics are interconnected as CDR encompasses the ethical use of digital technologies, including AI, ensuring that AI systems are designed, deployed, and utilized in ways that align with ethical principles and promote societal well-being (Sullivan & Wamba, 2022; Tóth et al., 2022).

To contribute to CDR research, we must be aware that different normative thoughts on ethics impact researchers' understanding of AI ethics and, thus, how they define responsible AI design, deployment, and usage (Lobschat et al., 2021). Siau and Wang (2020) stated that research on AI ethics concerns designing AI technologies according to ethical principles (ethics of AI) to enforce AI's ethical behavior (ethical AI). The perspective that AI must follow defined rules draws inspiration from deontological ethics as one normative ethical theory used by Weber (1988) and Kant (1797) and informs research to design AI in an ethically responsible manner based on duty-bound principles (e.g., Floridi et al., 2018; Jobin et al., 2019). However, AI ethics also considers how users' virtue ethics – representing humans' inherent moral understanding of right and wrong – impact ethical behavior and actions when interacting with AI (Heyder et al., 2023; Stolz et al., 2024). Researchers must be aware of how different ethical theories impact the understanding of responsible AI when transferring academic insights into practices (Brendel et al., 2021; Dignum, 2019).

The underlying ethical theory also influences how researchers perceive the moral agency of humans compared to AI (Coeckelbergh, 2020) – mainly when discussing who is responsible for undesired outcomes (Mueller, 2022). For instance, Allen et al. (2000, p. 252) argued that there is a "deep disagreement in ethical theory […] about what standards moral agents ought to follow." Those who follow the principle of utility – believing actions should achieve the maximum of aggregated good consequences (consequentialism)– argue that autonomous technologies should be programmed to reach this maximum from the bottom-up (Allen et al., 2000; Brendel et al., 2021), while others strive for bottom-up modeling of AI technologies based on virtue ethics theory making the "agents of AI (users, developers) … able to take responsibility for what they do with AI", by arguing that humans have to remain responsible since AI does not meet the traditional criteria of moral agency (Coeckelbergh, 2020, p. 2062).

Thus, different normative ethical theories impact our understanding of moral agency and AI responsibility from different perspectives (Dignum, 2019). To understand the contribution of AI ethics research and derive suggestions for responsible AI design, deployment, and usage following CDR (Mihale-Wilson et al., 2022), knowing which ethical theories ground the theoretical framing of AI ethics is essential. We overview the major normative ethical theories used in Information Systems (IS) research on AI ethics by focusing on deontological ethics, consequentialism, and virtue ethics (Bankins & Formosa, 2023; Bilal et al., 2021). Our research questions (RQs) are:

(RQ1): *What normative ethical theories are commonly used in AI ethics research?*
(RQ2): *What topics of AI ethics are studied the most within the IS field?*
(RQ3): *What are the calls for future investigations on AI ethics research?*

To answer these questions, we conduct a scoping review as it is appropriate to summarize academic discourse on a specific topic to identify the existing research landscape and potential gaps (Paré et al., 2015). We systematically review research on AI ethics within the IS field to determine which normative ethical theories focus on existing IS research and highlight areas that require further investigation.

Our contributions are twofold: *First*, we demonstrate how different normative ethical theories, implicitly or explicitly as theoretical frameworks, influence the interpretation of results and the discussion of findings in AI ethics research. The latter element is particularly relevant given the strong inclination towards deontological ethics, especially within the IS field, guiding researchers' suggestions for responsible AI design, deployment, and usage by combining ethics and human-computer interaction and technology application research as two disciplines of CDR (Lobschat et al., 2021). *Second*, we discuss the most studied topics in AI ethics research and offer recommendations for future research and practitioners to maximize the benefits of AI transformation.

2 Theoretical Background and Related Work

2.1 Normative Ethical Theories

Ethical theories guide how humans evaluate and form moral judgments (Beauchamp & Bowie, 1988). Within the domain of ethical studies, besides normative ethics, descriptive ethics, metaethics, and applied ethics can be analyzed (Brendel et al., 2021). We focus on normative ethics as it offers a systematic approach to determining moral duties and ethical conduct, providing prescriptive principles that dictate what actions are morally right or wrong (Brendel et al., 2021). Normative ethics provide guidance on how people should act. It encompasses various ethical theories, each proposing different principles for moral actions (Gewirth, 1960; Weaver & Trevino, 1994). Figure 1 provides essential normative ethical theories following Bankins and Formosa (2023):

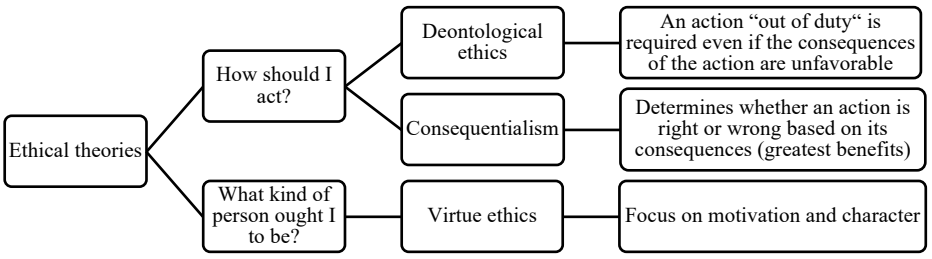

Figure 1. Normative Ethical Theories

Deontological ethics roots itself in duty and obligation, positing that specific actions are inherently moral or immoral, regardless of their outcomes (Gaus, 2001). Immanuel Kant introduced the concept of the "categorical imperative", which asserts that an action is morally right if it can be universalized without contradiction (Kant, 1797).

Consequentialism asserts that the morality of an action depends solely on its outcomes or consequences, determining the overall goodness or badness of results (Yu et al., 2018). The most well-known variant of consequentialism is utilitarianism, but other forms exist (e.g., ethical egoism), each with a different perspective on what outcomes are to be valued (Bentham, 2007; Mill, 2009). We focus on utilitarianism as it posits that the morality of an action is determined by the extent to which it maximizes overall happiness or pleasure (Bentham, 2007; Mill, 2009).

Virtue ethics, based on the works of Aristotle (Brown & Aristotle, 2009), emphasizes the importance of cultivating virtuous character traits, such as courage and honesty, and posits that moral actions stem from the inherent moral values of individuals (Eitel-Porter, 2021; Heyder et al., 2023). Vallor (2016) claims that, as emerging technologies grow more complex, virtue ethics provides a framework for moral decision-making amidst uncertainty, guiding us toward living well with these advancements.

2.2 Understanding of Moral Agency Forming Responsibility Perceptions

Central to normative ethical theories is the concept of moral agency that determines our perceptions of responsibility (Coeckelbergh, 2020; Passlack et al., 2023). Moral agency encompasses prevailing notions of "right" and "wrong" within a particular society and time, defined by a set of established norms and rules, both explicit and implicit (Behdadi & Munthe, 2020). It demonstrates the ability to choose the morally more responsible option (Champagne & Tonkens, 2023; Søvik, 2022). These moral values, norms, and belief systems constantly evolve and intertwine with historical social processes. Various social entities, such as religious institutions, criminal organizations, or cultural groups, craft their unique moral codes and adapt them over time, shaping distinct moralities. Hence, moral agency is required to attribute responsibility for specific actions (Constantinescu et al., 2021).

Central to moral agency is the concept of inner free will, which means the capacity to choose between actions, whereas AI does not have a moral agency in a traditional sense (Coeckelbergh, 2020). However, the increased autonomy of AI might lead people to perceive an Artificial Moral Agency (AMA) similar to a traditional human one and attribute responsibility to the AI, resulting in responsibility gaps (Behdadi & Munthe, 2020; Coeckelbergh, 2020; Giermindl et al., 2022). Thus, the CDR discourse rises (Lobschat et al., 2021) concerning ethical tensions from different responsibility perceptions of AI design, implementation, and usage.

Lobschat et al. (2021, p. 875) define CDR as "the set of shared values and norms guiding an organization's operations with respect to four main processes related to digital technology and data." While CDR focuses on the organizations' value system (corresponding to virtue ethics) and norms (corresponding to deontological ethics) in forming responsible or non-responsible actions, we aim to go one step before the organizations' CDR culture by investigating the researchers' perspectives that might influence CDR behavior as social context (Lobschat et al., 2021) by providing suggestions for responsible AI design, deployment, and usage that aligns human values and considers the ethical impact of AI based on different ethical theories (Dignum, 2019, p. 47).

2.3 Existing Literature Reviews on AI Ethics

Several systematic literature reviews summarize existing research on AI ethics within the IS field, each contributing unique insights and identifying prevailing ethical principles (Berente et al., 2021; Mirbabaie et al., 2022). However, a nuanced gap remains in understanding precisely which ethical theories have been used to better explain the underlying motivations for their use and its consequences on theoretical and practical contributions made by AI ethics researchers. For instance, Floridi et al. (2018) laid a significant foundation by introducing critical ethical principles for the establishment of a "Good AI Society" such as beneficence, non-maleficence, autonomy, justice/fairness, and accountability (p. 689). These principles ensure that organizations prevent technology from harming humankind. Jobin et al. (2019) expanded the scope by including grey literature in their systematic scoping review. They identified five primary ethical principles (transparency, justice and fairness, non-maleficence, responsibility, and privacy) and highlighted additional ones, broadening AI's moral discussion. Vainio-Pekka et al. (2023) further investigated the use of ethical principles specifically for explainable AI, reiterating the importance of the primary principles identified by previous research and pointing out the vagueness and lack of ethical knowledge as significant challenges. Khan et al. (2021) focused on ethical challenges and principles, echoing concerns about the vague nature of ethical guidelines and the prevalent lack of ethical knowledge. As Siau and Wang (2020) argued, research on AI ethics should not solely focus on exploring ethical principles for AI design but also on how to foster ethical behavior of the AI.

These reviews collectively emphasize the critical role of ethical principles in AI development, often referencing Jobin et al.'s (2019) global landscape of AI ethics guidelines and Floridi et al.'s (2018) AI4people principles as pivotal in motivating further research. The existing literature underscores the importance of these ethical principles in guiding AI development but often needs to delve deeper into the moral philosophy guiding their perspective on AI ethics. This gap highlights a crucial area for future research. Filling this gap could significantly advance the findings of reviews like those conducted by Brendel et al. (2021) and Heyder et al. (2023), which aim to understand the management of human-AI interaction from an ethical perspective by using different normative ethical theories for their literature analysis. Both works highlight how different ethical theories impact our understanding of moral or immoral behavior in the context of aiming to achieve ethical AI design, usage, and deployment.

3 Method

Our methodology builds on a systematic scoping review. A scoping review is appropriate to answer our research questions as it provides an overview of the nature of research on a specific topic to identify the breadth of coverage and potential research gaps (Paré et al., 2015). Unlike a systematic review, which synthesizes high-quality evidence to answer specific RQs and assess study quality, a scoping review maps all relevant literature irrespective of study design (Arksey & O'Malley, 2005). The systematic nature of following the steps of Templier and Paré (2018) increases the transparency:

1. Search strategy: Given our focus on IS research, we searched the following four databases: Association for Computing Machinery Digital Library (ACM), Association for Information Systems Electronic Library (AIS), Business Source Complete (BSC), and Institute of Electrical and Electronics Engineers Explore (IEEE). We searched for relevant literature by querying the title, abstracts, and keywords – if available – using the following search string: ("ethics" OR "morality" OR "ethical theory" OR "ethical family" OR "moral theory" OR "ethical philosophy" OR "moral philosophy" OR "normative ethics" OR "virtue ethics" OR "utilitarianism" OR "consequentialism" OR "deontology") AND ("Artificial Intelligence" OR "human-AI interaction"). We decided on these keywords after conducting initial exploratory research. The search conducted in September 2023 yielded 1234 overall hits. After eliminating 75 duplicates, we yielded 1159 unique hits. Table 1 illustrates the search process:

Table 1. Snapshot of Coding

Steps	ACM	AIS	BSC	IEEE	Total
Literature search	141	229	336	528	1234
Duplicates	1	1	46	27	75
Unique hits	140	228	290	501	1159
Format Screening	**424 excluded and 735 remaining articles**				
Non-scientific	32	1	218	85	336
Non-English	1	0	20	0	21
Incomplete	11	16	1	1	39
Stand-alone review	6	13	4	5	28
Relevance Screening	**674 excluded and 61 remaining articles**				
No ethical theory	10	37	5	47	99
Missing own perspective	9	26	1	7	43
No AI technology	0	0	1	13	14
No human-AI interaction	37	88	27	271	423
No dedicated contributions	7	26	6	56	95
Final sample	**27**	**21**	**7**	**6**	**61**
Top IS	**0**	**15**	**1**	**0**	**16**

2. Literature screening: We applied format criteria through format screening, eliminating non-scientific literature (e.g., poster presentations, thesis), non-English articles, incomplete articles (e.g., research-in-progress articles, extended abstract), and systematic and standalone literature reviews. Given the underlying goals of a scoping review, we did not restrict our search to a particular time range or specific quality criteria (e.g., rankings, peer-reviewed literature). This resulted in the elimination of 424 articles, and 735 articles remained for relevance screening. For relevance screening, we applied relevance criteria. We excluded articles that: i) cannot be related to ethical theories given their focus on other theoretical lenses (e.g., theory of planned behavior), ii) do not offer a meta-perspective on ethics from the paper's standpoint of view (e.g., ethics only appears in the context of an interviewees statement), iii) do not involve technologies that can be aligned to AI following the underlying definition of (Rai et al., 2019, p. iv) do not provide dedicated contributions regarding ethical theories (e.g., by solely focusing on one principle), do not focus on AI being used in several forms of human-AI interaction. 674 articles were excluded, and 61 articles remained as the final sample.

3. Classification of our final sample: Scoping reviews aim to be as comprehensive as possible in identifying primary studies and reviews (Arksey & O'Malley, 2005) and to achieve this, articles from journals of varying quality should be considered. To address the differences in quality between the articles of the final sample belonging to the core of IS' top journals and conferences, we followed a systematic approach to categorize the final sample into "top IS" (n=16) and other articles of the final sample not belonging to this category (n=45). Our list of "top IS" journals and conferences is based on several sources[1]: we involved IS conferences and journals being part of the VHB JOURQUAL3 Ranking, the Financial Times Top 50 journal list ("FT50"), the "basket of the 11" following Lowry et al. (2013), the list of "pure IS" having a rating of 0.8 or higher as outlined by Walstrom and Hardgrave (2001), and the framework of assimilating journals being ranked A+, A, or B as proposed by Levy and Ellis (2006).

4. Qualitative analysis: The 61 articles of our final sample were analyzed through manual coding. First, we structured the samples' meta-information (e.g., regarding the research domain, method, and the ranking of the publication outlet) to provide a descriptive overview of the different article types within the final sample. We then opted for qualitative analysis to structure the content of relevant articles. We used the theoretical background to structure the findings mainly deductively (Mayring, 2014, pp. 95-98). To increase the quality of the coding process and redefine the coding schema, we used 32 articles for parallel training purposes of the coders based on the coding schema given in Table 2. Utilizing Cohen's Kappa, as Cohen (1960) outlined, we measured interrater agreement, obtaining scores of α=0.76 in the initial round and α=0.82 in the subsequent round. These results signify a substantial agreement between coders, which aligns with McHugh's interpretation (2012). To enhance coding reliability, we conducted multiple discussion rounds among coders. Table 1 presents a snapshot of the coding results; Table 2 shows the final sample related to different AI research areas.

4 Results

Regarding our **first research questions** on what normative ethical theories are most commonly used in AI ethics research, our review highlights a clear focus on deontological ethics, with 37.70% involving specific contracts to be obtained when designing or managing AI and 20.97% following specific ethical AI principles. Only two articles discussed all three common views on normative ethics: Allen et al. (2000) stated that understanding artificial moral agents varies with ethical theories. Vanhée and Borit (2022, p. 623) argued that using "classic ethical theories, such as deontic, utilitarian, and virtue ethics" as moral exemplars can impact ethical behavior. Hence, our final sample does not present virtue ethics and consequentialism (each 4.84%), demonstrating a research gap, especially as "moral values need to guide technologists in their research and to help decision-makers regulate the use that is made of the possibilities that AI presents from their deeper understanding of the possibilities and the consequences" (Gómez de Ágreda, 2020, pp. 3-4).

[1] https://ethicaltheories.github.io/

Table 2. Snapshot of Coding

Category	Explanation	Codes	Coding Rules	Examples
Normative ethical theory	States whose view on normative ethics is specified or assigned.	Deontology Consequentialism Virtue ethics Contractarianism Other	We use the definitions provided in Section 2. If the theory is not explicitly stated but discussed on a meta-level that aligns with a particular theory, it may be coded "implicit"; see category "presence of ethical theory").	**Code for virtue ethics:** "Virtue ethics focuses on moral character development rather than duties, rules, or the consequences of actions to decide what is right or wrong." (Bilal et al., 2021, p. 2) **Codes for deontology & consequentialism:** "The two approaches considered are utilitarianism, on the one hand, and Kant's use of the `categorical imperative' on the other." (Allen et al., 2000, p. 252)
Presence of ethical theory	Identifies whether the ethical theory is explicitly mentioned or implicitly applied in the research	Explicit Implicit Absent	If the research does not directly mention an ethical theory but its principles are recognizable through the discussion, scenarios, or examples provided, we code it as "implicit." If an ethical theory is directly mentioned, named, or defined in the research, we code it as "explicit."	**Code for explicit (virtue):** "…virtue ethics, understood as an approach to normative ethics that emphasizes moral character in contrast to approaches that emphasize duties and rules (deontology) or consequences of actions consequentialism)" (Vanhée & Borit, 2022, p. 619) **Code for implicit (deontological):** "We are assuming that ethical and moral reasoning by humans is fundamentally computational and does not involve some special human-only force or quality." (Seo & Thorson, 2022, p. 1077)

Table 3. Overview of the final sample related to different research categories

Category	Explanation	Example	All (62)	Top IS* (16)
Regulating AI ethics	Articles in this category have an institutional or regulatory focus on AI ethics.	Almeida et al., 2020*; Auld et al., 2022; Colmenarejo et al., 2022; Cooper et al., 2022; Fabiano, 2019; Ha, 2022; Henderson, 2019; Lim & Kwon, 2021; Polyviou & Zamani, 2022*; Seppälä et al., 2021*; Tidjon & Khomh, 2022; Unver, 2023; Walke et al., 2023; Weber, 2020; Westerstrand, 2023*; Wright, 2020	16	4
Designing AI ethics	Articles in this category mainly deal with the design of ethical AI, i.e., technical or design-oriented guidelines, as well as actual technical solutions or computational models.	Benner et al., 2021*; Bilal et al., 2021; Chaput et al., 2021; Cook, 2023; Eicher et al., 2018; Emdad et al., 2023*; Gerdes, 2022; Hooker & Kim, 2018; Jantunen et al., 2021; Ong, 2021; Schlimbach & Khosrawi-Rad, 2022*; Susser, 2019; Vakkuri et al., 2021; Vanderelst & Winfield, 2018; Vanhée & Borit, 2022	15	3
Researching AI ethics	Articles in this category contribute to research on AI ethics.	Aslan et al., 2022*; Dyrkolbotn et al., 2018; Estrada, 2018; Greene et al., 2019*; Hawkins & Mittelstadt, 2023; Mirbabaie et al., 2022*; Seymour, 2018; Siapka, 2022; Siau & Wang, 2020; Vainio-Pekka et al., 2023; Zhang et al., 2021	11	3
Managing AI ethics	Articles in this category mainly contain managerial approaches to dealing with AI ethics.	Agbese et al., 2023; Akbari Ghatar et al., 2023*; Corvite et al., 2023; Figueras et al., 2022; Gómez de Ágreda, 2020; Mäntymäki et al., 2023*; Mayer et al., 2021*; Minkkinen & Mäntymäki, 2023; Rismani & Moon, 2023; Wang et al., 2020*	10	4
Educating AI ethics	Articles in this category mainly introduce approaches that include ethical aspects when teaching and training AI (e.g., in higher education institutions, such as computer sciences).	Forsyth et al., 2021; Kowch, 2019; McDonald & Pan, 2020; Raji et al., 2021; Schneider et al., 2023*	5	1
Theorizing AI ethics	Articles in this category mainly develop new theories by taking a philosophical lens.	Allen et al., 2000; Burema et al., 2023; Seo & Thorson, 2022*; Terzis, 2020	4	1

Considering the impact of ethical theory on how researchers address the concept of moral agency, future work on AI ethics must involve different ethical theories. As Allen et al. (2000, p. 253) stated: "There is, then, a clear sense of 'morally good' which a utilitarian can apply to an agent's actions irrespective of how the agent decided upon that action (consequentialism). Not so for Kant [deontology]. According to Kant, for an action to be morally good, the action must, as he puts it, be done out of respect for the categorical imperative. Here we follow one account of the categorical imperative and assume that 'acting out of respect for it' means simply that the agent acted as it did because it determined that the action in question was consistent with the categorical imperative. In this sense of the term, an action cannot be morally good unless the agent, in fact, reasoned in certain fairly complex ways."

In the "top IS" articles, 62.50% focus on applied ethics, 12.50% on consequentialism, and 6.25% each on virtue ethics and deontology. Applied ethics often underlies a deontological view that is either implicit (18.75%) or unclaimed (absent, 18.75%). While 32.73% of all 61 articles in the final sample explicitly claimed the use of ethical theory, 62.50% of the "top IS" articles explicitly outlined the underlying ethical theory. Future researchers might clarify their understanding of AI ethics, especially since there can be "…tensions between practitioners' interpretation of ethical principles in their work and ethos tensions. In this vein, we argue that understanding the tensions that can occur in practice and how they are tackled is key to studying ethics in practice. Understanding how AI practitioners perceive and apply ethical principles is necessary for practical ethics to contribute toward an empirically grounded, Responsible AI." (Figueras et al., 2022, p. 1).

Regarding our **second research question** on AI ethics within IS research, we see that AI ethics research, design, and management are deeply interconnected with AI regulation, education, and theorization, as detailed below.

When *researching AI ethics*, researchers often apply a societal lens, delving into themes like social justice (e.g., Greene et al., 2019), fair play (e.g., Estrada, 2018), feminism (e.g., Siapka, 2022) corresponding to metaethics, or explainable AI (e.g., Vainio-Pekka et al., 2023) and AI safety (e.g., Zhang et al., 2021) corresponding to applied ethics. In doing so, researchers favor investigations that favor breadth over depth, which might be due to the difficulties in aligning different perceptions of what is right or wrong depending on different normative ethical theories that are used within different investigational settings (Seo & Thorson, 2022). The complexity of AI ethics emerges as a recurring theme. Mirbabaie et al. (2022) and Siau and Wang (2020) point out the lack of clear definitions and conceptualizations of AI ethics and the absence of a general theory to explain its ethical dimensions from a normative point of view.

Designing ethical AI revolves around the central role of ethical principles following deontological ethics. Benner et al. (2021) and Schlimbach and Khosrawi-Rad (2022) highlight the importance of these principles, while Emdad et al. (2023), Hooker and Kim (2018), Vakkuri et al. (2021), and Vanhée and Borit (2022) focus on guidelines and frameworks. The importance of experimentation and learning ethical behavior through participatory approaches (involving human users' virtue ethics) is also emphasized, as discussed by Eicher et al. (2018), Chaput et al. (2021), and Gerdes (2022).

This complements research on nudging (i.e., guiding individuals' behavior toward a beneficial choice for themselves or society) (Mirbabaie et al., 2022).

Managing AI ethics involves addressing different perceptions of moral agency (in particular, AMA) and, thus, how to attribute responsibility (e.g., Dyrkolbotn et al., 2018; Figueras et al., 2022; Gómez de Ágreda, 2020; Rismani & Moon, 2023; Wang et al., 2020). Research on organizational practices following the organizational stakeholders' perceptions of what is right or wrong (virtue ethics), such as the work of Agbese et al. (2023) and Akbari Ghatar et al. (2023), reveals varying practices for evaluating fairness and the difficulty of reaching a shared definition of fairness in AI ethics. This perspective is further underlined by the work of Figueras et al. (2022) and Mayer et al. (2021), who analyze the perceptions of responsible AI practices in organizations. Their findings align with the challenges of a missing shared understanding of AI ethics. A few articles specifically address managing responsible AI by aligning organizational aims of AI implementation, such as increasing working efficiency (corresponding to consequentialism) with CDR norms (corresponding to deontological ethics) and workforces and leadership values (corresponding to virtue ethics) as seen in the work of Figueras et al. (2022), Rismani and Moon (2023), and Wang et al. (2020). These studies contribute to a broader understanding of how organizations manage AI ethics.

Regulating AI ethics involves addressing its societal impact and exploring potential regulatory and legal approaches. Our review highlights a growing call for training and educating AI governance topics (Mäntymäki et al., 2023; Minkkinen & Mäntymäki, 2023). In doing so, structured approaches with defined ethical principles (corresponding to deontological ethics) might guide AI technologies' ethical use and development within organizations and at a broader societal level (corresponding to humans' virtue ethics). The emphasis on governance reflects an understanding that AI's ethical considerations are not limited to design principles (deontological ethics) but require active and ongoing management and oversight to ensure responsible use (virtue ethics). Therefore, appropriate training and education are required.

When theorizing AI ethics, sector-based perspectives are vital. Burema et al. (2023) emphasize that understanding AI systems within specific sectoral cultures is crucial, given that perceptions of responsibilities are different depending on the ethical understanding of different cultures (specifically, virtue ethics). Terzis (2020) addresses the abstract nature of AI ethics discussions and the "illusion of agreement" among stakeholders due to varying interpretations of ethical behavior based on their virtue ethics.

5 Discussion

5.1 Theoretical Implications

Our review outlines two theoretical implicatoins (corresponding to our third RQ).

We need a more fundamental discussion on how our ethical understanding impacts our perceptions of responsibilities: Mayer et al. (2021) present a framework that shows how companies can encourage their employees to implement ethical AI and

to take responsibility for specific actions, using different combinations of implementation measures and mechanisms of AI ethics policies: employee engagement, organizational embedding, practical support, and assurance processes. Mirbabaie et al. (2022) further claim that IS research lacks a general theory and fundamental research on the ethical dimensions in the complex AI context guiding our understanding of responsibilities. In addition, Vakkuri et al. (2021) go straight to the heart of the issue and call for an in-depth examination of the term ethical AI itself. The need for a definition and an examination of the terminology is also evident as some authors (e.g., Agbese et al., 2023; Akbari Ghatar et al., 2023; Corvite et al., 2023) use the term ethical AI without defining it, which might increase readers' perceptions on an AI that might be accountable resulting in responsibility gaps (Coeckelbergh, 2020). Cooper et al. (2022) also mention that future work should increasingly address the question of which actors should take responsibility in an algorithmic society and to what extent.

We need more diverse and empirical research on how different value systems impact AI ethics principles: Empirical investigations are needed to compare AI ethics principles grounded in the value systems of various industries and company types, utilizing globally diverse sample sizes (Agbese et al., 2023; Jantunen et al., 2021; Mayer et al., 2021; Seppälä et al., 2021). For instance, publicly listed companies may be subject to more pressure from stakeholders (Mayer et al., 2021) and, therefore, implement AI ethics measures earlier. Alternatively, a geographical focus can result in a lack of generalizability as value systems differ between countries and their underlying cultures, forming their norms on what might be right or wrong guiding AI adoption guidelines (Agbese et al., 2023; Seppälä et al., 2021). Cross-sectoral (Minkkinen & Mäntymäki, 2023) and comparative studies can provide insights into industries, sectors, and regional differences, taking into account highly regulated areas such as medicine or less ethically sensitive areas such as manufacturing (Mäntymäki et al., 2023). Future studies should also compare the implications of AI ethics between the private and public sectors (Figueras et al., 2022). According to Corvite et al. (2023), developing close-ended surveys might contribute to marginalized groups or vulnerable individuals (McDonald & Pan, 2020). As AI technologies and the associated ethical considerations are developing so rapidly, researchers must continuously monitor and account for this development (Burema et al., 2023), wherefore new contexts should continuously be added (Forsyth et al., 2021; Lim & Kwon, 2021) and longitudinal studies should be implemented (Seppälä et al., 2021). Schneider et al. (2023) argue that empirical research would benefit the discourse on ethical AI.

5.2 Limitations and Future Work

Our scoping review has potential limitations. First, the search strategy did not include backward and forward searches, which might have limited the comprehensiveness of the literature review by excluding relevant articles not indexed in the four databases searched. Additionally, our search string did not encompass all potential keywords related to AI, such as "Machine Learning" (ML), which may have resulted in missing some relevant studies. Lastly, we focused on specific databases for keyword search.

Due to the interdisciplinary nature of the AI ethics field (Islam & Greenwood, 2023), future work might integrate further literature, such as from the Ethical Philosophy field. Future research might extend our insights by investigating AI ethics in the context of applied, descriptive, and meta-ethics, given the clear dependence between the theoretical justification of AI ethics and the interpretation of results.

5.3 Practical Implications

Our review highlights a gap between theory and practice (Tidjon & Khomh, 2022; Vainio-Pekka et al., 2023; Weber, 2020), which is critical as practitioners are those implementing AI solutions (Vainio-Pekka et al., 2023). Following Mirbabaie et al. (2022), the knowledge gained in the theoretical works should create a scientific foundation as AI influences society and individuals. According to the authors, the influence of NGOs and the media on AI use should also be investigated in practice. Best practices can bridge the gap from theory to practice, supporting "the creation of new methods, tools, frameworks, and guidelines to implement ethical AI principles into practice" (Seppälä et al., 2021, p. 15). This transfer is crucial since Mayer et al. (2021) underline that little is known about AI ethics in practice.

One possibility for practitioners might be the orientation on the most often used and cited ethical principles for AI. Aslan et al. (2022, p. 2) stated: "[…] the most common definitions and frameworks are AI4People (Floridi et al., 2018), the OECD Principles of AI (OECD 2019), and the Ethics Guidelines for Trustworthy AI ('High-level expert group on Artificial Intelligence' (HLEG; AI, 2019))." Besides the HLEG and OECD principles, the authors of our final sample also mentioned the AI Act of the European Commission (e.g., Tidjon & Khomh, 2022; Walke et al., 2023; Westerstrand, 2023), the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (e.g., Figueras et al., 2022; Lim & Kwon, 2021; Ong, 2021), and the General Data Protection Regulation (GDPR; e.g., Almeida et al., 2020; Fabiano, 2019; Siapka, 2022).

Practitioners can also integrate AI ethics into curricula since AI-related knowledge and experiences might contribute to more responsible usage of AI (Weber et al., 2023). For instance, Henderson (2019) argues that ethical discourse should be expanded and integrated into computer science curricula. Additionally, McDonald and Pan (2020) support the integration of intersectional AI thinking into students' curricula.

6 Conclusion

Our scoping review on AI ethics research synthesizes IS literature to investigate what normative ethical theories are most commonly used, what topics of AI ethics are studied most and calls for future work in this field. The review demonstrates that IS research lacks the integration of several views on normative ethics, resulting in assessing ethics mainly from the AI's perspective by integrating ethical principles into AI's design, rather than considering how a human user's virtue equally impacts human-AI implications (e.g., through their data input and the way they critically reflect and use AI outputs).

References

Agbese, M., Mohanani, R., Khan, A., & Abrahamsson, P. (2023), 'Implementing AI Ethics: Making Sense of the Ethical Requirements', *in* 'Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering', pp. 62–71.

Akbari Ghatar, P., Pappas, I., & Vassilakopoulou, P. (2023), 'Practices for Responsible AI: Findings from Interviews with Experts', *in* 'AMCIS 2023 Proceedings'.

Allen, C., Varner, G., & Zinser, J. (2000), 'Prolegomena to any future artificial moral agent', *Journal of Experimental & Theoretical Artificial Intelligence*, **12**(3), pp. 251–261.

Almeida, P., Santos, C., & Farias, J. S. (2020), 'Artificial Intelligence Regulation: A Meta-Framework for Formulation and Governance', *in* 'Proceedings of the Hawaii International Conference on System Sciences (HICCS) 2020'.

Arksey, H., & O'Malley, L. (2005), 'Scoping studies: Towards a methodological framework', *International Journal of Social Research Methodology*, **8**(1), pp. 19–32.

Aslan, A., Greve, M., & Lembcke, T.-B. (2022), 'Let´s Do Our Bit: How Information Systems Research Can Contribute to Ethical Artificial Intelligence', *in* 'AMCIS 2022 Proceedings'.

Auld, G., Casovan, A., Clarke, A., & Faveri, B. (2022), 'Governing AI through ethical standards: Learning from the experiences of other private governance initiatives', *Journal of European Public Policy*, **29**(11), pp. 1822–1844.

Baird, A., & Maruping, L. (2021), 'The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts', *Management Information Systems Quarterly*, **45**(1), pp. 315–341.

Bankins, S., & Formosa, P. (2023), 'The Ethical Implications of Artificial Intelligence (AI) for Meaningful Work', *Journal of Business Ethics*, **4**, pp. 1–16.

Beauchamp, T. L., & Bowie, N. E. (1988), *Ethical Theory and Business* (Subsequent Edition), Prentice Hall.

Behdadi, D., & Munthe, C. (2020), 'A Normative Approach to Artificial Moral Agency', *Minds and Machines*, **30**(2), pp. 195–218.

Benner, D., Schöbel, S., & Janson, A. (2021), 'It is only for your own good, or is it? Ethical Considerations for Designing Ethically Conscious Persuasive Information Systems', *in* 'AMCIS 2021 Proceedings'.

Bentham, J. (2007), *An Introduction to the Principles of Morals and Legislation*, Dover Publications Inc.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021), 'Managing Artificial Intelligence', *MIS Quarterly*, **45**, pp. 1433–1450.

Bilal, A., Wingreen, S., Sharma, R., & Jahanbin, P. (2021), 'Trust Development in Artificial Intelligence-based Emerging Technologies: Rise of Technomoral Virtues and Data Ethics', *in* 'ACIS 2021 Proceedings'.

Brendel, A. B., Mirbabaie, M., Lembcke, T.-B., & Hofeditz, L. (2021), 'Ethical Management of Artificial Intelligence', *Sustainability*, **13**(4), Article 4.

Brown, L., & Aristotle. (2009), *The Nicomachean Ethics* (D. Ross, Trans.; New Edition).

Burema, D., Debowski-Weimann, N., von Janowski, A., Grabowski, J., Maftei, M., Jacobs, M., van der Smagt, P., & Benbouzid, D. (2023), 'A sector-based approach to AI ethics: Understanding ethical issues of AI-related incidents within their sectoral context', *in* 'Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society', pp. 705–714.

Champagne, M., & Tonkens, R. (2023), 'A Comparative Defense of Self-initiated Prospective Moral Answerability for Autonomous Robot harm', *Science and Engineering Ethics*, **29**(4), p. 27.

Chaput, R., Duval, J., Boissier, O., Guillermin, M., & Hassas, S. (2021), 'A Multi-Agent Approach to Combine Reasoning and Learning for an Ethical Behavior', *in* 'Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society', pp. 13–23.

Coeckelbergh, M. (2020), *AI Ethics*, The MIT Press.

Colmenarejo, A. B., Nannini, L., Rieger, A., Scott, K. M., Zhao, X., Patro, G. K., Kasneci, G., & Kinder-Kurlanda, K. (2022), 'Fairness in Agreement With European Values: An Interdisciplinary Perspective on AI Regulation', *in* 'Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*'*, pp. 107–118.

Constantinescu, M., Voinea, C., Uszkai, R., & Vică, C. (2021), 'Understanding responsibility in Responsible AI. Dianoetic virtues and the hard problem of context', *Ethics and Information Technology*, **23**, pp. 803–814.

Cook, T. (2023), 'Robust Artificial Moral Agents and Metanormativity', *in* 'Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society', pp. 162–169.

Cooper, A. F., Moss, E., Laufer, B., & Nissenbaum, H. (2022), 'Accountability in an Algorithmic Society: Relationality, Responsibility, and Robustness in Machine Learning', *in* '2022 ACM Conference on Fairness, Accountability, and Transparency', pp. 864–876.

Corvite, S., Roemmich, K., Rosenberg, T. I., & Andalibi, N. (2023), 'Data Subjects' Perspectives on Emotion Artificial Intelligence Use in the Workplace: A Relational Ethics Lens', *in* 'Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1)', pp. 124:1-124:38.

Dignum, V. (2019), *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*, Springer Verlag.

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., … Wright, R. (2023), 'Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy', *International Journal of Information Management*, **71**, 102642.

Dyrkolbotn, S., Pedersen, T., & Slavkovik, M. (2018), 'On the Distinction between Implicit and Explicit Ethical Agency', *in* 'Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society', pp. 74–80.

Eicher, B., Polepeddi, L., & Goel, A. (2018), 'Jill Watson Doesn't Care if You're Pregnant: Grounding AI Ethics in Empirical Studies', *in* 'Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society', pp. 88–94.

Eitel-Porter, R. (2021), 'Beyond the promise: Implementing ethical AI', *AI and Ethics*, **1**(1), pp. 73–80.

Emdad, F. B., Ho, S., Ravuri, B., & Hussain, S. (2023), 'Towards A Unified Utilitarian Ethics Framework for Healthcare Artificial Intelligence', *in* 'AMCIS 2023 Proceedings'.

Estrada, D. (2018), 'Value Alignment, Fair Play, and the Rights of Service Robots', *in* 'Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society', pp. 102–107.

Fabiano, N. (2019), 'Robotics, intelligent systems, ethics and data protection: The new challenge', *in* 'Proceedings of the 2nd International Conference on Applications of Intelligent Systems', pp. 1–5.

Figueras, C., Verhagen, H., & Pargman, T. C. (2022), 'Exploring tensions in Responsible AI in practice. An interview study on AI practices in and for Swedish public organizations', *Scandinavian Journal of Information Systems*, **34**(2).

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', *Minds and Machines*, **28**(4), pp. 689–707.

Forsyth, S., Dalton, B., Foster, E. H., Walsh, B., Smilack, J., & Yeh, T. (2021), 'Imagine a More Ethical AI: Using Stories to Develop Teens' Awareness and Understanding of Artificial Intelligence and its Societal Impacts', *in* '2021 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)', pp. 1–2.

Gaus, G. (2001), 'What is Deontology? Part One: Orthodox Views', *The Journal of Value Inquiry*, **35**, pp. 27–42.

Gerdes, A. (2022), 'A participatory data-centric approach to AI Ethics by Design', *Applied Artificial Intelligence*, **36**(1), p. 2009222.

Gewirth, A. (1960), 'Meta-Ethics and Normative Ethics', *Mind*, **69**(274), pp. 187–205.

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', *European Journal of Information Systems*, **31**(3), pp. 410–435.

Gómez de Ágreda, Á. (2020), 'Ethics of autonomous weapons systems and its applicability to any AI systems', *Telecommunications Policy*, **44**(6), p. 101953.

Greene, D., Hoffmann, A. L., & Stark, L. (2019), 'Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning', *in* 'Proceedings of the Hawaii International Conference on System Sciences 2019 (HICSS)'.

Ha, Y. J. (2022), 'South Korean Public Value Coproduction Towards 'AI for Humanity': A Synergy of Sociocultural Norms and Multistakeholder Deliberation in Bridging the Design and Implementation of National AI Ethics Guidelines', *in* 'Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency', pp. 267–277.

Hawkins, W., & Mittelstadt, B. (2023), 'The ethical ambiguity of AI data enrichment: Measuring gaps in research ethics norms and practices', *in* '2023 ACM Conference on Fairness, Accountability, and Transparency', pp. 261–270.

Henderson, T. (2019), 'Teaching Data Ethics: We're going to ethics the heck out of this', *in* 'Proceedings of the 3rd Conference on Computing Education Practice', pp. 1–4.

Heyder, T., Passlack, N., & Posegga, O. (2023), 'Ethical management of human-AI interaction: Theory development review', *The Journal of Strategic Information Systems*, **32**(3), p. 101772.

HLEG; AI. (2019), 'Ethics guidelines for trustworthy AI', https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai, [last visited June 28, 2024].

Hooker, J. N., & Kim, T. W. N. (2018), 'Toward Non-Intuition-Based Machine and Artificial Intelligence Ethics: A Deontological Approach Based on Modal Logic', *in* 'Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society', pp. 130–136.

Islam, G., & Greenwood, M. (2023), 'Ethical Research in Business Ethics', *Journal of Business Ethics*, **182**(1), pp. 1–5.

Jantunen, M., Halme, E., Vakkuri, V., Kemell, K.-K., Rousi, R., Mikkonen, T., Duc, A. N., & Abrahamsson, P. (2021), 'Building a Maturity Model for Developing Ethically Aligned AI Systems', *Selected Papers of the IRIS*, Issue Nr 12 (2021).

Jobin, A., Ienca, M., & Vayena, E. (2019), 'The global landscape of AI ethics guidelines', *Nature Machine Intelligence*, **1**(9), Article 9.

Kant, I. (1797), *Die Metaphysik der Sitten*, Nicolovius.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020), 'Algorithms at Work: The New Contested Terrain of Control', *Academy of Management Annals*, **14**(1), pp. 366–410.

Khan, A. A., Badshah, S., Liang, P., Khan, B., Waseem, M., Niazi, M., & Akbar, M. A. (2021), 'Ethics of AI: A Systematic Literature Review of Principles and Challenges (arXiv:2109.07906)', *arXiv*.

Kowch, E. G. (2019), 'Ethics to Prepare Educators for Professional Service Robots in Classrooms', *in* '2019 International Joint Conference on Information, Media and Engineering (IJCIME)', pp. 478–484.

Lim, J. H., & Kwon, H. Y. (2021), 'A Study on the Modeling of Major Factors for the Principles of AI Ethics', *in* 'DG.O2021: The 22nd Annual International Conference on Digital Government Research', pp. 208–218.

Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., & Wirtz, J. (2021), 'Corporate digital responsibility', *Journal of Business Research*, **122**, pp. 875–888.

Lowry, P. B., Moody, G. D., Gaskin, J., Galletta, D. F., Humpherys, S., 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?', *Management Information Systems Quarterly*, **37**(4), 993–1012.

Mäntymäki, M., Minkkinen, M., Zimmer, M., Birkstedt, T., & Viljanen, M. (2023), 'Designing an AI governance framework: From research-based premises to meta-requirements', *in* 'ECIS 2023 Proceedings'.

Mayer, A.-S., Haimerl, A., Strich, F., & Fiedler, M. (2021), 'How corporations encourage the implementation of AI ethics', *in* 'ECIS 2021 Proceedings'.

Mayring, P. (2014), *Qualitative content analysis—Theoretical foundation, basic procedures and software solution*, Klagenfurt.

McDonald, N., & Pan, S. (2020), 'Intersectional AI: A Study of How Information Science Students Think about Ethics and Their Impact', *in* 'Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2)', pp. 147:1-147:19.

McHugh, M. L. (2012), 'Interrater reliability: The kappa statistic', *Biochemia Medica*, **22**(3), pp. 276–282.

Mihale-Wilson, C., Hinz, O., van der Aalst, W., & Weinhardt, C. (2022), 'Corporate Digital Responsibility', *Business & Information Systems Engineering*, **64**(2), pp. 127–132.

Mill, J. S. (2009), *Utilitarismus*, Felix Meiner Verlag.

Minkkinen, M., & Mäntymäki, M. (2023), 'The Institutional Logics Underpinning Organizational AI Governance Practices', *in* 'Proceedings of the 14th Scandinavian Conference on Information Systems'.

Mirbabaie, M., Brendel, A., & Hofeditz, L. (2022), 'Ethics and AI in Information Systems Research', *Communications of the Association for Information Systems*, **50**.

Mueller, B. (2022), 'Corporate Digital Responsibility', *Business & Information Systems Engineering: The International Journal of WIRTSCHAFTSINFORMATIK*, **64**(5), pp. 689–700.

Ong, D. C. (2021), 'An Ethical Framework for Guiding the Development of Affectively-Aware Artificial Intelligence (arXiv:2107.13734)', *arXiv*.

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015), 'Synthesizing information systems knowledge: A typology of literature reviews', *Information & Management*, **52**(2), pp. 183–199.

Passlack, N., Hammerschmidt, T., Klemm, F., & Posegga, O. (2023), 'How Human-AI Collaboration Affects Attribution of Responsibility for Failure and Success', *in* 'ICIS 2023 Proceedings'.

Polyviou, A., & Zamani, E. (2022), 'Towards Europe's AI Strategy: Desires & Realities', *in* 'ECIS 2022 Proceedings'.

Rai, A., Constantinides, P., & Sarker, S. (2019), 'Editor's Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids', *Management Information Systems Quarterly*, **43**(1), pp. iii–ix.

Raji, I. D., Scheuerman, M. K., & Amironesei, R. (2021), 'You Can't Sit With Us: Exclusionary Pedagogy in AI Ethics Education', *in* 'Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency', pp. 515–525.

Rismani, S., & Moon, Aj. (2023), 'What does it mean to be a responsible AI practitioner: An ontology of roles and skills', *in* 'Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society', pp. 584–595.

Schlimbach, R., & Khosrawi-Rad, B. (2022), 'Towards Ethical Design Features for Pedagogical Conversational Agents', *in* 'AMCIS 2022 Proceedings'.

Schneider, J., Meske, C., & Bikic, A. (2023), 'How individuals can shape AI through data—An AI literacy and morality perspective', *in* 'ECIS 2023 Proceedings'.

Seo, H., & Thorson, S. (2022), 'Computation, Rule Following, and Ethics in AIs', *in* 'Proceedings of the Hawaii International Conference on System Sciences 2022 (HICSS)'.

Seppälä, A., Birkstedt, T., & Mäntymäki, M. (2021), 'From Ethical AI Principles to Governed AI', *in* 'ICIS 2021 Proceedings'.

Seymour, M. (2018), 'Artificial Intelligence Is No Match for Human Stupidity: Ethical Reflections on Avatars and Agents', *in* 'ACIS 2018 Proceedings'.

Siapka, A. (2022), 'Towards a Feminist Metaethics of AI', *in* 'Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society', pp. 665–674.

Siau, K., & Wang, W. (2020), 'Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI', *Journal of Database Management (JDM)*, **31**(2), pp. 74–87.

Søvik, A. O. (2022), 'What overarching ethical principle should a superintelligent AI follow?', *AI & Society*, **37**(4), pp. 1505–1518.

Stolz, K., Hammerschmidt, T., & Posegga, O. (2024), 'How Conversational Agents Influence Purchase Decisions of Online Fashion Shoppers toward Sustainable Consumption: Exploring Nudges for Green Decision-Making', *in* 'Proceedings of the Hawaii International Conference on System Sciences 2024 (HICSS)'.

Sullivan, Y. W., & Wamba, S. F. (2022), 'Moral Judgments in the Age of Artificial Intelligence', *Journal of Business Ethics*, **178**(4), pp. 917–943.

Susser, D. (2019), 'Invisible Influence: Artificial Intelligence and the Ethics of Adaptive Choice Architectures', *in* 'Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society', p. 1.

Templier, M., & Paré, G. (2018), 'Transparency in literature reviews: An assessment of reporting practices across review types and genres in top IS journals', *European Journal of Information Systems*, **27**(5), pp. 503–550.

Terzis, P. (2020), 'Onward for the freedom of others: Marching beyond the AI ethics', *in* 'Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency', pp. 220–229.

Tidjon, L. N., & Khomh, F. (2022), 'The Different Faces of AI Ethics Across the World: A Principle-Implementation Gap Analysis (arXiv:2206.03225)', *arXiv*.

Tóth, Z., Caruana, R., Gruber, T., & Loebbecke, C. (2022), 'The Dawn of the AI Robots: Towards a New Framework of AI Robot Accountability', *Journal of Business Ethics*, **178**(4), pp. 895–916.

Unver, M. B. (2023), 'Rebuilding "ethics" to govern AI: How to re-set the boundaries for the legal sector?', *in* 'Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law', pp. 306–315.

Vainio-Pekka, H., Agbese, M. O., Jantunen, M., Vakkuri, V., Mikkonen, T., Rousi, R., & Abrahamsson, P. (2023), 'The Role of Explainable AI in the Research Field of AI Ethics', *ACM Transactions on Interactive Intelligent Systems*.

Vakkuri, V., Kemell, K.-K., Jantunen, M., Halme, E., & Abrahamsson, P. (2021), 'ECCOLA—A method for implementing ethically aligned AI systems', *Journal of Systems and Software*, **182**, p. 111067.

Vallor, S. (2016), *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*, Oxford University Press.

Vanderelst, D., & Winfield, A. (2018), 'The Dark Side of Ethical Robots', *in* 'Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society', pp. 317–322.

Vanhée, L., & Borit, M. (2022), 'Viewpoint: Ethical By Designer - How to Grow Ethical Designers of Artificial Intelligence', *Journal of Artificial Intelligence Research*, **73**, pp. 619–631.

Walke, F., Bennek, L., & Winkler, T. (2023), 'AI in Government: A Study on Explainability of High-Risk AI-Systems in Law Enforcement & Police Service', *in* 'Wirtschaftsinformatik 2023 Proceedings'.

Walstrom, K. A., & Hardgrave, B. C. (2001), 'Forums for information systems scholars: III', *Information & Management*, **39**(2), 117–124.

Wang, Y., Xiong, M., & Olya, H. (2020), 'Toward an Understanding of Responsible Artificial Intelligence Practices', *in* 'Proceedings of the Hawaii International Conference on System Sciences 2020 (HICSS)'.

Weaver, G. R., & Trevino, L. K. (1994), 'Normative and Empirical Business Ethics: Separation, Marriage of Convenience, or Marriage of Necessity?', *Business Ethics Quarterly*, **4**(2), pp. 129–143.

Weber, M. (1988), *Gesammelte Aufsätze zur Wissenschaftslehre*, Mohr Siebeck GmbH & Co. KG.

Weber, P., Pinski, M., & Baum, L. (2023), 'Toward an Objective Measurement of AI Literacy', *in* 'PACIS 2023 Proceedings'.

Weber, R. H. (2020), 'Socio-ethical values and legal rules on automated platforms: The quest for a symbiotic relationship', *Computer Law & Security Review*, **36**, p. 105380.

Westerstrand, S. (2023), 'Ethics in the intersection of AI and democracy: The AIDEM Framework', *in* 'ECIS 2023 Proceedings'.

Wright, S. A. (2020), 'AI in the Law: Towards Assessing Ethical Risks', *in* '2020 IEEE International Conference on Big Data (Big Data)', pp. 2160–2169.

Yu, H., Shen, Z., Miao, C., Leung, C., Lesser, V. R., & Yang, Q. (2018), 'Building Ethics into Artificial Intelligence (arXiv:1812.02953)', *arXiv*.

Zhang, B., Anderljung, M., Kahn, L., Dreksler, N., Horowitz, M. C., & Dafoe, A. (2021), 'Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers (arXiv:2105.02117)', *arXiv*.