Received: 10 June 2023 | Accepted: 7 December 2023

DOI: 10.1111/basr.12336

ORIGINAL ARTICLE

BSR Business and Society Review | WILEY

What ethics can say on artificial intelligence: Insights from a systematic literature review

Francesco Vincenzo Giarmoleo[1] | **Ignacio Ferrero**[1] | **Marta Rocchi**[2] | **Massimiliano Matteo Pellegrini**[3]

[1]School of Economics and Business, University of Navarra, Pamplona, Spain

[2]DCU Business School, Irish Institute of Digital Business, Dublin City University, Glasnevin, Ireland

[3]Management and Law Department, School of Economics, Università degli Studi di Roma Tor Vergata, Rome, Italy

Correspondence
Francesco Vincenzo Giarmoleo, School of Economics and Business, University of Navarra, Campus Universitario, Pamplona 31009, Spain.
Email: fgiarmoleo@unav.es

Funding information
No grants or specific funding was received to assist with the preparation of this manuscript.

[Correction added on 12 February 2024, after first online publication: In the byline, Massimiliano Matteo Pellegrini's middle name and ORCID has been added.]

Abstract

The abundance of literature on ethical concerns regarding artificial intelligence (AI) highlights the need to systematize, integrate, and categorize existing efforts through a systematic literature review. The article aims to investigate prevalent concerns, proposed solutions, and prominent ethical approaches within the field. Considering 309 articles from the beginning of the publications in this field up until December 2021, this systematic literature review clarifies what the ethical concerns regarding AI are, and it charts them into two groups: (i) ethical concerns that arise from the design of AI and (ii) ethical concerns that arise from human–AI interactions. The analysis of the obtained sample highlights the most recurrent ethical concerns. Finally, it exposes the main proposals of the literature to handle the ethical concerns according to the main ethical approaches. It interprets the findings to lay the foundations for future research on the ethics of AI.

KEYWORDS

AI, artificial intelligence, ethics

1 | INTRODUCTION

The debate on the ethics of artificial intelligence (AI) is lively and multifaceted. Some authors identify ethical issues regarding the design, use, and deployment of AI-based systems and their impact on business and society (Coeckelbergh, 2021; Martin, 2019; Tollon, 2021). Others discuss what moral status should be granted to machines (Awad et al., 2019; Smith & Vickers, 2021) and how to deal with the "responsibility gap," when no clear party is held responsible for the actions performed by an AI system (Orr & Davis, 2020). Other contributions discuss the challenges presented by human–machine interaction (Losbichler & Lehner, 2021; Miller, 2019), privacy protection (Guha et al., 2021; McStay, 2020), or the impact on specific fields, such as business strategies. An example of the latter is Callanan et al. (2021), who focus on data mining and automated prediction strategies. This growing literature can be attributed to the continuously increasing pervasiveness of AI and 4.0 revolution instruments in business and in the everyday life of society at large (Schwab, 2016; World Economic Forum, 2023). While these technologies and their improvements are beneficial from many perspectives, they invariably also draw our attention to the problems and concerns they may give rise to.

Given the number and variety of themes and perspectives in AI ethics, a systematization of this vast body of knowledge is a particularly desirable contribution. Consequently, many valuable efforts have been directed towards obtaining a net systematization of the literature. However, most of them tend to focus on categorizing ethical issues related to a specific field (Borges et al., 2021; Hunkenschroer & Luetge, 2022; Morley et al., 2020) or on guiding principles (Jobin et al., 2019; Khan et al., 2021). But these contributions fail to propose solutions to handle these issues, or when they do, they omit any analysis of the ethical approaches that underpin them. And this analysis is important because it represents a bridge between academic research and business practices in utilizing AI technology for the betterment of society and human well-being. In addition, by understanding the underlying moral philosophies, we can have more meaningful and coherent discussions about AI ethics, and it can help us identify and address weaknesses in existing approaches, rendering them more effective in dealing with AI's complex ethical challenges. For these reasons, this article presents a systematic literature review to investigate prevalent concerns, proposed solutions, and prominent ethical approaches that aim to strengthen the ways in which ethical concerns in AI ethics are addressed.

Hence, the work described in this article fulfills a threefold purpose. First, it identifies the most relevant ethical concerns in the literature of AI. Second, it describes the main proposals and solutions in existing academic literature to handle these concerns. Third, it investigates the ethical approaches that these solutions are based on. To achieve these aims, the article follows the methodology of a systematic literature review that analyzes 309 articles from 1986 until December 31, 2021. The article is structured around the stages suggested by Rowley and Slack (2004): (i) giving a basic definition of the topic; (ii) clarifying why the subject is of interest; (iii) expounding what research has already been undertaken on the topic; and (iv) providing a clear summary of the research opportunities and objectives that emerge from the literature review.

Following this structure, Section 2 briefly presents the definition and importance of ethics in AI. Then, it discusses existing reviews on the ethics of AI identifying research gaps. This is followed by an analysis of the protocol, the search, the criteria, and the quality assessment adopted in this article. Section 3 presents the quantitative and thematic analysis of the 309 academic articles of the sample, describing the solutions found in the literature to manage AI-related ethical concerns. It also examines—as a novelty in the literature—the ethical

approaches supporting these proposals. Finally, the conclusion offers a critical assessment of the findings and identifies new research paths.

2 | CHARACTERIZING AI ETHICS

2.1 | The ethical dimension of AI and previous contributions

According to the definition given by the latest revision of the European Union (EU) AI act, AI is "a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments" (European Parliament, 2023, p. 137). This definition is in line with the OECD's definition[1] and underscores several crucial aspects of AI-based systems: their goal-oriented essence, varying levels of autonomy, the breadth of generated outputs, and their potential influence across diverse sectors.

The academic literature highlights two main reasons why an ethical analysis of AI-based technologies is important. First, technology is never an unstoppable force of nature but always a human product (Harambam et al., 2018), and therefore, it is a vehicle of moral intentions due to the cultural context, pushing the user to a specific act (Benanti, 2019) and must be subjected to moral discernment (Rosenberger & Verbeek, 2015). Nevertheless, in the AI field, there is an attitude of technological determinism (Greene et al., 2019). In this context, contributions from ethicists would be crucial to prevent that "the task of understanding the nuances emerging from analytics is just left to statisticians" (Grodzinsky, 2017, pp. 227–228). According to Gabriel (2020), the challenge of integrating the ethical perspective within AI-based technologies design is twofold and regards both the technical ability to encode it formally and what principles should be encoded. Second, as noticed by Dignum (2018), ethics and AI are related at several levels. There is ethics *by* design, which is "the technical/algorithmic integration of ethical reasoning capabilities as part of the behavior of an artificial autonomous system" (p. 2). Another is ethics *in* AI design, involving the regulatory and engineering techniques that enable the examination and appraisal of the ethical consequences stemming from the integration or replacement of conventional social structures by AI systems. Finally, there is ethics *for* AI design, which is "the codes of conduct, standards and certification processes that ensure the integrity of developers and users as they research, design, construct, employ and manage artificial intelligent systems" (Dignum, 2018, p. 2).

The abundant literature on ethical concerns in AI illustrates the importance given to the topic by academic research. Several literature reviews have been published since 2016 with the aim to systematize existing literature on AI from different perspectives. Table 1 offers an overview of these publications, highlighting the scope chosen by the authors when addressing their systematic literature review and the total number of articles considered.

Only a few contributions listed in Table 1 adopt both a general approach to AI ethics and the objective to map the main ethical concerns or principles (Ashok et al., 2022; Jobin et al., 2019; Khan et al., 2021; Meek et al., 2016). Additionally, the existing literature does not provide a clear exposition of the solutions proposed to face these concerns nor does it explain the ethical frameworks adopted to solve ethical concerns related to AI. We think this latter analysis is very important for the reasons outlined above.

This article systematizes the literature from 1986 to 2021 in the AI ethics field, with the specific aim to map ethical concerns (Research Question 1) and solutions (Research Question 2),

TABLE 1 The focus of the previous literature reviews

Article	Scope of the systematic literature review	Number of articles considered	General or specific subfield	Aim to map AI's ethical challenges or principles
Meek et al. (2016)	Assessing AI's development stage, future potential, and ethical risks in its implementation.	/	General	Yes
Reijers et al. (2018)	Analyzing and discussing methods for practicing ethics within the realm of research and innovation (R&I).	73	Specific	No
Carvalho et al. (2019)	Presenting a systematic literature review of machine learning methods applied to predictive maintenance.	54	Specific	No
Jobin et al. (2019)	Analyzing principles and guidelines to determine whether there exists a global consensus on ethical requisites, technical standards, and optimal practices.	84	General	Yes
Hagerty and Rubinov (2019)	Assessing the societal influence of AI across five worldwide regions.	800 academic journal articles and monographs.	General	No
Burr et al. (2020)	Identifying the central issues across diverse social domains related to digital wellbeing, which encompasses the impact of digital technologies on the notion of a fulfilling human life.	179	General	No
Di Vaio et al. (2020)	Inquiring whether AI possesses the capability to influence production and consumption patterns, aligning with sustainable resource management objectives as outlined in the sustainable development goals (SDGs) of the UN's 2030 agenda.	73	General	No
Glikson and Woolley (2020)	Conducting empirical cross-disciplinary investigations to identify the factors influencing human trust in AI across various sectors and disciplines.	150	General	No
Morley et al. (2020)	Examining the debate surrounding the ethical considerations of AI within the domain of healthcare.	147	Specific	Yes

TABLE 1 (Continued)

Article	Scope of the systematic literature review	Number of articles considered	General or specific subfield	Aim to map AI's ethical challenges or principles
Borges et al. (2021)	Analyzing the integration of AI into organizational strategy considering existing approaches, benefits, challenges, and opportunities, and potential research directions.	41	Specific	Yes
Khan et al. (2021)	Investigating the agreement and challenging factors on AI principles.	27	General	Yes
Möllmann et al. (2021)	Examination of AI's benefits and ethical concerns in healthcare, specifically digital health.	50	Specific	Yes
Toorajipour et al. (2021)	Analyzing the prevalent and potential AI techniques in supply chain management (SCM), the current AI-enhanced SCM subfields, and the subfields with high potential for AI improvement.	64	Specific	No
Votto et al. (2021)	Examining the application of AI in human resource management.	33	Specific	No
Zuiderwijk et al. (2021)	Analzying the application of AI in public governance.	26	Specific	Yes
Ashok et al. (2022)	Highlighting the impact of digital ethics implications on society.	59	General	Yes
Bao and Xie (2022)	Providing a systematic review on the application of AI in animal farming.	131	Specific	No
Gomes et al. (2022)	Individuating AI-based methods that possess the potential to automate business processes and aid companies' decision-making procedures.	21	Specific	No
Hunkenschroer and Luetge (2022)	Mapping the ethical opportunities, risks, ambiguities and proposals to the AI technology in recruiting and selection process.	51	Specific	Yes

and characterize which ethical approach supports the solutions offered (Research Question 3). While a few previous publications have dealt with Research Question 1, we substantially expand the scope and impact of this systematization by categorizing solutions and corresponding ethical approaches. As we show in the analysis that follows, and make evident throughout the article, the identification and categorization of the proposals are key to building a bridge between academic research and business practice when it comes to the use of AI-based technology in a way that promotes human flourishing and the good of society.

2.2 | Research design and methodology

The research questions (RQ) that guided the study are as follows:

- RQ1: What are the ethical concerns in AI explored by academic literature?
- RQ2: What are the recurrent solutions to manage the concerns regarding AI ethics?
- RQ3: What are the ethical approaches underpinning the solutions in AI ethics?

The research questions were instrumental in formulating the keywords and the search strategy, the inclusion and exclusion criteria, and the quality assessment. The systematic literature review adopted the method of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2009). This article aims to contribute to the field by integrating and augmenting the current literature and providing new insights. In particular, it offers the following:

1. A map of the ethical concerns (answer to RQ1).

We classify them according to two main categories: (i) concerns that arise from the design of AI and (ii) concerns that arise from human–AI interactions. Compared to the previous classifications, this article extends the analysis by identifying 21 ethical concerns, and it provides the frequency of occurrence of these concerns in the literature, mapping the main lines of debate for each.

2. A map of the solutions to handle these ethical concerns (answer to RQ2).

This contribution represents a novelty within the existing literature. Among other contributions, Jobin et al. (2019) and Khan et al. (2021) focused on principles, while this article extends and integrates their perspective, through a map of proposals to handle the ethical concerns.

3. A map of the main ethical approaches explicitly or implicitly underpinning proposed solutions (answer to RQ3).

This final part of the research aims to set out the leading ethical approaches in the formulation of proposals to handle ethical concerns regarding AI. A detailed description of the leading ethical approaches in AI is essential for guiding ethical AI development, fostering interdisciplinary collaboration, and ensuring that proposed solutions are robust and practical. Through this analysis, the article also seeks to identify gaps and meaningful paths for future research.

The search was performed in EBSCO and ABI Inform for articles published up to and including December 31, 2021, containing "ethic*" AND "artificial intelligence" OR "AI" in the title, abstract and keywords. The field of AI is characterized by including several perspectives and without a standard evaluation criteria (Collins et al., 2021). Therefore, we select "AI" or "Artificial Intelligence" as search terms based on the following reasons. First, it is the exact string used in other systematic literature reviews, such as in Ashok et al. (2022). Second, the term "artificial intelligence" is commonly used in the jargon to describe AI-based systems (Guo et al., 2019). Third, the choice is due to the article's focus, which aims to analyze the various sectors where AI-based technologies are used, without overrepresenting some subfields at the expense of others. Indeed, Gomes et al. (2022) noted that "artificial intelligence" is an umbrella that includes other subfields, such as machine learning and expert systems, or different technologies: Mustak et al. (2021) in marketing, Chahal and Byrne (2020) and Mutasa et al. (2020) in medicine, Zawacki-Richter et al. (2019) in education, and Reim et al. (2020) in business model innovation.

Finally, the research endeavored to capture the ethical dimension of AI, analyzing all the results that refer to "ethic," "ethics," and "ethical." The choice is due to the following reasons. First, the expression "ethics of artificial intelligence" is used in the literature as an expression that indicates the ethical aspects of both robots and machines (in terms of design) and of human behavior in the interaction with AI-based technologies (Guo et al., 2019). Second, this choice was successfully made also by previous systematic literature reviews, such as Morley et al. (2020). Finally, this choice allows to extend the approach of previous publications, for example, adopting a more inclusive search than Ashok et al. (2022). The research opened the most extensive timeline available in the databases, which was up to and including December 31, 2021. AI's history starts from the half of the last century, so choosing a specific date as a starting point for the literature review would have been arbitrary. The first article identified is from 1986.

2.3 | The inclusion and exclusion criteria, the quality assessment and the data extraction

The search was conducted in EBSCO in Business Source Ultimate, E-Journals, MLA Directory of Periodicals, MLA International Bibliography, and Philosopher's Index up until 2021. In the first phase of identification, we obtained 2014 articles. We then refined the search by limiting it to peer-reviewed results, which resulted in 1118 articles, and finally, we filtered for contributions in academic journals only, eliminating conference proceedings, commentaries, editorials, back materials, book reviews, news, roundtables, and interviews, which narrowed it down to 232 results. The same search was repeated in ABI Inform, obtaining 1,687 results. After filtering for peer-reviewed publications and for contributions in academic journals, we ended up with a final list of 359 articles. In the second phase of screening, we compared the articles obtained through ABI Inform and EBSCO by eliminating duplicated records. Consequently, we obtained a list of 524 results. We selected only the publications in English, as per other systematic literature reviews, such as Busalim (2016), which meant we were left with 499 articles. Next, we moved on to manually screening the dataset with the goal of identifying and excluding any non-scientific articles that may not have been filtered out automatically earlier, reducing the sample further to 399 articles.

Finally, after having read all the abstracts, we excluded some articles that were not coherent with the purpose of our work, because they were neither focused on AI (32) nor on AI's ethical

concerns (38). Finally, we eliminated 20 contributions that appeared just as abstracts. As a result, we were left with 309 articles that constitute the final sample.

Figure 1 depicts the phases of identification, screening, and inclusion foreseen by the PRISMA protocol:

The methodology for the thematic analysis of the articles is as follows:

First, there was an abstract review phase, during which two of the authors meticulously examined the abstracts. They determined the inclusion or exclusion of each article based on its congruence with the aims of this article. Instances of differing opinions were resolved through discussion, and to ensure objectivity in the paper selection process, the effectiveness of this approach was verified through Krippendorf's alpha coefficient, which had to be more than 0.80, signifying robust inter-reliability (Krippendorff, 2019).

After obtaining the final sample, the remaining articles were subjected to a thematic classification and analysis by two of the authors, who for each article isolated the main themes and

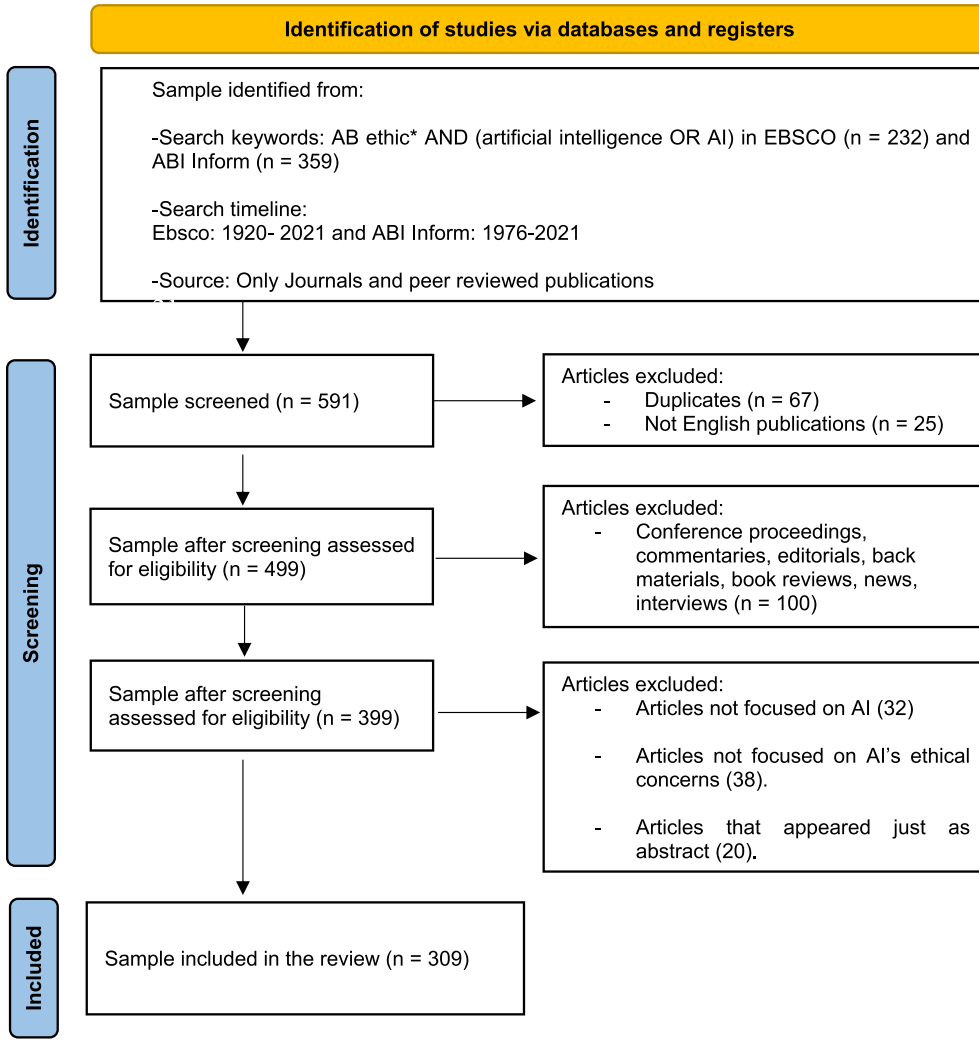

FIGURE 1 The selection process

contributions. They meticulously recorded on an excel spreadsheet the relevant details for each paper, both of a quantitative nature, e.g., article references, publication year, and journal outlet, and of a qualitative nature, according to our multi-layered structure, e.g., main ethical concerns, proposals to address them, and supporting ethical foundations. The framework that was adopted made provision for any paper to have multiple focuses on different ethical concerns. For example, a paper focused on discrimination could explore both algorithm fairness and unequal participation. Again, the individual classifications were cross-checked by the two researchers and a similar level of agreement regarding the qualitative dimension was found (again Krippendorf's alpha coefficient of more than 0.80). The final framework was finally reviewed collectively by the whole team of authors to assure coherence and rigor. The resulting framework found 21 emerging ethical concerns divided into two groups, namely, concerns that arise from the design and concerns that arise from human–AI interaction (details about this classification are given in Section 3.1). The framework also classifies proposals according to their manner in which they deal with the ethical concerns based on an act- or agent-centered ethical orientation (details about this classification are given in Section 3.2).

3 | RESULTS OF THE THEMATIC ANALYSIS

Looking at the databases analyzed in this study, we can offer some general insights about the knowledge structure of the field of AI and ethics. Table 2 describes the composition of the sample in relation to the year of publication.

Due to the rapid development of AI technology during the last few years, 246 out of the 309 articles (80%) have been published between 2019 and 2021. The journal that publishes the most is *Ethics and Information Technology*, with 26 articles. It is followed by *Philosophy & Technology* with 22 articles, *Computer Law & Security Review* with 13 articles, *Big Data & Society* with 12 articles, and *Journal of Information, Communication & Ethics in Society* with 9 articles. Furthermore, it is noteworthy that a significant majority of the articles (70.22%) have opted for a thematic research methodology. In contrast, a relatively smaller proportion (5.83%) has chosen a quantitative approach. A mixed methodology, blending both qualitative and quantitative methods, has been embraced by 23.95% of the articles. The articles focusing on business can be categorized as follows: the largest portion pertains to Administration and Management (42.7%), followed by Marketing (21.5%), Accounting (12%), Human Resources (9.5%), Logistics (7.1%), Customer Service (4.8%), and Finance (2.4%).

TABLE 2 Number of academic publications per group-years

Years	Number of academic publications
1986–1990	1
1991–1995	1
1996–2000	1
2001–2005	4
2006–2010	2
2011–2015	14
2016–2021	286

We offer an overview of the topic, and we report the findings of some of the articles of the sample included in each category (concerns, proposal, and ethical foundations). Considering the size of the dataset used in this systematic literature review, it would be impossible to offer insights from all the articles contained in each category. For this reason, we selected the most relevant articles for each category according to two criteria: first, of the degree to which the content of the paper matched the issue analyzed and, second, the relevance of the paper measured by the number of citations the paper received from articles published in the same year, and then a weighted citations index determined by dividing the number of citations a paper receives by the total. This metric is designed to account for the time dimension, as older papers tend to accumulate more citations than newer ones. Similar to a GPS program that can pinpoint someone's geographic location, we think this way of presenting the literature can help readers of AI content "find their compass" in the large corpus of AI ethics literature.

3.1 | The analysis of the ethical concerns

The thematic analysis aims to report on the focus of each article. The main ethical concerns raised by AI in literature can be classified into two groups: (i) concerns that arise from the design of AI and (ii) concerns that arise from human–AI interactions. This categorization has been used mirroring theoretical classifications already affirmed in the field of AI and in the literature. First, the definition of AI as adopted in the EU AI Act emphasizes both these aspects. Indeed, there is focus on design elements and varying autonomy levels ("A machine-based system that is designed to operate with varying levels of autonomy," i.e., design concerns), while it also concentrates on the consequences of system outputs on the physical or virtual environment ("Generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments," i.e., human–AI interaction concerns). Second, similar classifications have already been used in the literature, as the distinction drawn and introduced by Meek et al. (2016) between the ethical issues that arise from the AI technology itself and issues that arise from the effects of AI on living beings. In addition, Guo et al. (2019) in defining the ethics of AI make a distinction between robot ethics and machine ethics. The former pertains to humans involved in envisioning and designing AI systems during the initial phases of their development (again, design concerns), while the latter delves into the consequences of AI systems, encompassing both intended and unintended effects (again, human–AI interaction concerns). Finally, this classification aligns with the scope of the article in that it is adaptable to various scenarios and fields, capturing how diverse groups, such as developers, decision-makers, and the general public are impacted by AI.

3.1.1 | Ethical concerns about the design of AI

This section reports on the emerging ethical concerns regarding *how* AI is designed and *who* designs it, based on their occurrence within the literature review sample.

Algorithm and data

More than 20% of the contributions are centered on the ethical dimensions of algorithms and data. This theme can be further categorized into two main subthemes: data bias and algorithm fairness, and algorithm opacity.

Data bias and algorithm fairness (12.3%). This category encompasses two distinct research streams. The first one delves into the social consequences of data bias and algorithm fairness. Helberger et al. (2020) present findings from a survey of the Dutch adult population, revealing that AI-driven automated decision-making systems are perceived as fairer than human decision-makers by many respondents. Weber (2020) highlights the deployment of AI on automated platforms and the need for a legal framework that safeguards socio-ethical values and fundamental rights. The article underscores a trust-based approach rooted in human values to balance AI's market and technology-driven use with a multi-stakeholder perspective. Astobiza et al. (2021) introduces the concept of "algorithmocracy," discussing algorithms' transformative impact on society and the imperative to imbue AI systems with ethics to achieve the 17 Sustainable Development Goals. Schiff et al. (2021) address the disparity between ethical principles and the practical implementation of AI, elucidating how barriers impede the translation of ethical aspirations into practice.

The second research stream focuses on practical methodologies to mitigate bias. Janssen et al. (2020) propose a data governance framework that promotes stewardship of data, processes, and algorithms, alongside transparency, shared ownership, and self-sovereign identities. Vakkuri et al. (2021) present a method for implementing AI ethics through an iterative research approach involving researchers and practitioners.

Algorithm opacity (7.8%). This subtheme gives rise to two distinct strands of research. The first one explores the necessity for regulations and indications for policymakers to ensure the responsible development of AI. For instance, Hongladarom (2021) scrutinizes Thailand's National AI Ethics Guideline, shedding light on Thailand's efforts to modernize while upholding its traditional values. Chiao (2019) investigates the ethical ramifications of using AI within the realm of criminal justice. The study assesses whether algorithms can enhance existing modes of decision-making in this context, presenting cautious optimism alongside potential concerns.

The second strand entails practical methodologies to address algorithmic opacity within specific domains. For example, Brown et al. (2021) introduce an auditing framework designed to evaluate the ethical impact of algorithms. This framework encompasses stakeholder interests, pertinent ethical metrics, and a matrix that links these metrics to stakeholder concerns. Stefan and Carutasu (2020) propose an enhanced machine learning model that incorporates the concept of explicability to facilitate ethical AI-based decisions. They present a framework for extending machine learning models with an explanatory layer, enabling users to comprehend the decision-making process behind AI-generated outcomes. Thiebes et al. (2021) center their work on the notion of trustworthy AI (TAI), which rests on five foundational principles: beneficence, non-maleficence, autonomy, justice, and explicability. The authors outline a research framework for TAI that is data-driven, and they suggest avenues for research and practical application utilizing distributed ledger technology.

Balancing AI's risks

This category constitutes more than 16% of the articles and focuses on addressing the potential risks associated with AI systems. Given the ubiquity of AI technologies, these articles explore the implications of AI risks across various contexts linked to design and unpredictability, military purposes, emergency procedures, and AI takeover.

Design faults and unpredictability (9.2%). A key concern within this group revolves around design faults, in particular new processes to enhance the safety of AI systems. For instance, Siafakas (2021) investigates innovative procedures for AI scientists, while Donia and Shaw

(2021) examine the role that co-designing plays in tackling ethical challenges posed by AI in healthcare. They assess the effectiveness of co-designing in managing these challenges and highlight potential pitfalls.

Military and security purposes (3.8%). This group concerns the deployment of AI for military applications. Taddeo et al. (2021) present an ethical framework for AI use in defense, emphasizing transparency, human responsibility, and reliable AI systems. Mathew and Mathew (2021) study the ethical dilemma of deploying autonomous weapon systems in warfare and the significance of human oversight in preventing civilian casualties. Another research line explores normative and social considerations linked to this issue. Sari and Celik (2021) provide a legal evaluation of AI-based lethal weapon system attacks, addressing accountability and responsibility.

Emergency procedures (1.9%). This theme revolves around preparing for emergencies in AI systems, specifically focusing on strategies, ethical considerations, and practical measures to ensure swift and effective responses in unforeseen circumstances. One research line discusses alternative approaches to conventional emergency interventions. Arnold and Scheutz (2018) propose continuous self-evaluation and testing as alternatives to "big red button" interventions, aiming to prevent system errors and risks in AI systems. Another line focuses on new procedures to tackle emerging issues in specific contexts. Winkle et al. (2018) conducted a broad analysis of accidents stemming from poor visibility, providing real-world test scenarios for the safe development and validation of automated vehicles under challenging conditions.

AI takeover (1.7%). This group represents articles envisioning scenarios where advanced AI systems attain autonomy and control. One research line examines the implications of delegating decision-making to AI. Rakowski et al. (2021) discuss AI's dual role as a tool for human emancipation and a potential risk, touching on social control, security, democracy, and values. Kamishima et al. (2018) propose a capability-effectual approach, which combines the capability approach with the effectual process model from entrepreneurship research, to ethical AI entrepreneurship for risk management and responsible innovation. Another research line engages in broader discussions about AI's future, incorporating a precautionary principle into AI technologies. Simanowski et al. (2019) explore uncertainties tied to artificial consciousness and challenges in governing AI evolution. McLean et al. (2021) provide a systematic review of risks associated with Artificial General Intelligence (AGI): an AI system able to simulate human intelligence across a wide range of activities, underlining the need to address limitations in AGI risk research.

Threats to human institutions and life

This group comprises 11% of the articles and centers on risks stemming from AI systems designed with malicious intent or that can end up in a threat to human life. It can be divided into two key themes: threats to law and democracy, and transhumanism.

Threats to law and democratic values (10.6%). This theme underscores the ethical dilemmas AI poses to democratic values and human rights. One subset of research revolves around methodologies and frameworks for assessing AI's impact on fundamental rights. Mantelero and Esposito (2021) propose a comprehensive approach to data protection impact assessments, considering the collective impact on fundamental rights and societal values. Aizenberg and van

den Hoven (2020) address challenges in designing AI systems that respect human rights and values. They discuss how efficient algorithms in decision-making can lead to unjust outcomes and emphasize bridging the gap between abstract values and design requirements.

Another research line is concerned with AI's societal impact. De Almeida et al. (2021) present a conceptual framework for AI regulation, integrating diverse aspects of modern public policy-making to uphold societal values like fairness and freedom. Reisach (2021) discusses the responsibility of social media platforms in the context of societal and political manipulation. The article addresses concerns about the influence of machine learning techniques on citizen decisions, as seen in events like the Brexit referendum. Formosa et al. (2021) develop a framework for analyzing ethical issues in cybersecurity contexts such as penetration testing, ransomware, and system administration. Wamba et al. (2021) focus on preparing for a "good AI society," offering a bibliometric review and research agenda. They identify 10 social impact domains related to AI research and present 136 research questions that address AI-driven societal changes.

Transhumanism (0.5%). This category highlights the inherent uncertainty of transhumanism, which seeks to surpass biological limitations by merging humans with AI technologies. This advancement raises inquiries regarding the distinction between humans and machines, the potential outcomes of this integration, and the ethical reflections concerning improved human capabilities via AI enhancements. The ethical concern revolves around the exploration of this unfamiliar domain of human–AI fusion, which raises critical questions about identity, self-governance, parity, and the plausible advantages and drawbacks linked to surpassing biological restrictions. Nicola and Dalessio (2019) reference Pandya (2019)'s statement about the potential consequences of developing neuromorphic chips capable of processing human sensory data and responding without explicit programming. These advancements could give rise to beings with superhuman abilities, often referred to as "cyborgs," and could potentially lead to the emergence of a new species that may not necessarily prioritize humanity's best interests. Belk (2021) explores ethical concerns around service robotics and AI technologies, including transhumanism, and emphasizes the importance of research and policy considerations. In particular, the author examines the merger of humans and AI to enhance capabilities and potentially achieve immortality, by uploading human consciousness into AI, for example. The resulting disparity in opportunities between people with abundant material resources and those from disadvantaged situations could exacerbate existing inequalities.

Uniformity in the AI field

This group of concerns represents 2% of the sample and highlights two central issues: Western centrality and cultural difference, and unequal participation.

Western centrality and cultural differences (1.3%). This concern addresses the intersection of cultural diversity and ethical dimensions within the field of AI. Some articles emphasize the need to infuse AI ethics and governance with diverse socio-cultural perspectives. Segun (2021) underscores the influence of global socio-cultural factors on ethics and their role in shaping AI ethics, particularly in the African context. ÓhÉigeartaigh et al. (2020) tackle barriers to cross-cultural collaboration in AI ethics and governance, identifying challenges to cooperation between different regions and cultures in AI development. Pak-Hang (2020) explores the challenge posed by cultural disparities in global AI ethics and governance, examining the human rights approach to AI governance. Additionally, some research delves into non-Western philosophical traditions

to offer alternative ethical frameworks for AI development (Carman & Rosman, 2021; Shakir et al., 2020).

Unequal participation (0.9%). This concern focuses on unequal participation in the AI field. Oleksy et al. (2012) concentrate on gender-related issues within literature on specific Information and Communication Technologies (ICTs). The authors discuss gender concerns associated with these technologies and propose that certain ICTs, for example, those powered by AI, could positively impact gender power dynamics and promote gender balance in labor markets. Quinn (2021) emphasizes the issue of representation gaps, where both minority groups and women are inadequately represented or misrepresented in data.

3.1.2 | Ethical concerns about human–AI interactions

This section reports on the ethical concerns associated with the interaction between humans and AI as seen in the literature review. The four macro-categories extracted from the analysis of the literature are presented below.

Building a human–AI environment

This category encompasses nearly 17% of the articles and addresses the overall imperative of establishing a harmonious coexistence between humans and machines, and the key concerns that gives rise to this need.

Impact on business (10.1%). This group is concerned with the ethical implications of AI's influence on business models and practices in general but also on specific business practices. Dwivedi et al. (2021) explore AI and social media's impact on consumer behavior and business operations, addressing ethical concerns and offering research propositions. Munoko et al. (2020) examine the ethical implications of AI in auditing, discussing practical and social issues linked to AI's application in auditing. Fritts and Cabrera (2021) address concerns about dehumanization in hiring processes due to AI recruitment algorithms, advocating for the maintenance of substantive employee-employer relationships. Others tackle general implications of AI's impact on business. Cavallone and Palumbo (2020) review the implications of Industry 4.0, AI, and digitalization in healthcare, highlighting their potential to enhance care quality while addressing governance, ethics, and management challenges. Sako (2020) emphasizes that professionals need hybrid capabilities encompassing AI expertise and domain specialization to contribute to understanding how AI shapes future professional work and the evolving nature of skills and roles in businesses as a result.

Impact on jobs (5.7%). This subset underscores the potential consequences of AI on employment and the workforce. Some papers examine the broader influence of AI within the context of Industry 4.0 and propose corresponding strategies. Melé (2021) analyzes ethical risks posed by Industry 4.0, exploring effects on employment, inequality, human treatment, safety, and meaningful work from a Catholic Social Teaching standpoint. Yam and Skorburg (2021) highlight human rights implications of hiring algorithms, presenting a framework to evaluate and audit the impact of such algorithms on the human rights of job applicants. Other articles explore the possibility of AI replacing humans in specific domains. Chen et al. (2020), for example, study public attitudes towards AI replacing human doctors, revealing differing opinions and

emphasizing the need to consider public sentiments in promoting medical AI. Watson et al. (2021) investigate the evolving role of senior leaders in an AI-driven landscape, identifying essential capabilities for navigating AI's influence, including digital proficiency, ethics, and adaptability. Alderman (2021) discusses AI's potential to widen gender gaps in the accounting profession and advocates for supporting women's technological literacy and emotional intelligence skills.

Accessible AI (1.1%). This group assesses the ethical dimensions of AI's impact on accessibility, with a focus on integrating vulnerable communities. Siau and Wang (2020) reports the risk of the lack of accessibility for the elderly. Montes and Goertzel (2019) notice that the field of AI is presently controlled by a small group of centralized mega-corporations, functioning as an oligopoly, and their primary focus is directed towards the concerns of their stakeholders. They propose a decentralized and democratized AI market using distributed ledger technology that has as its aim, equitable AI and AGI (Artificial General Intelligence) development.

Privacy protection

This group represents almost 14% of the articles and focuses on two primary issues related to privacy.

Privacy threats to citizens (10.5%). This subset underscores the need for global regulations and governance mechanisms to ensure privacy in the context of AI technologies. Hickman and Petrin (2021) examine the influence of the EU's Trustworthy AI Guidelines, emphasizing the necessity for better alignment with existing company laws and governance principles while addressing AI's potential risks to privacy. Lilkov (2021) evaluates the EU's comprehensive draft Regulation on AI, analyzing provisions, potential implementation challenges, and argues for a flexible legal framework in AI regulation. Other publications address specific challenges related to AI privacy. Hamilton et al. (2021) address privacy concerns tied to Voice Activated Personal Assistants (VAPAs) like Alexa and Siri. They review legal issues, propose policy guideline categorization, and suggest a user-informed rating system. McStay (2020) identifies a weak consensus on privacy among stakeholders and highlights the need for societal agreements on privacy principles in emotional AI practices, referring to the integration of AI technologies into everyday objects and practices capable of detecting and responding to human emotions. Williams et al. (2020) discuss the shift from word of mouth (WOM) to algorithmic word of mouth (aWOM) as a result of AI advancements, emphasizing aWOM's potential dominance in influencing tourist decisions, ethical deployment, and privacy concerns.

Privacy threats to customers (3.3%). This research line addresses AI's impact on marketing and customer relations. Du and Xie (2021) explore ethical challenges in AI-enabled consumer markets, discussing biases, privacy, cybersecurity, and individual autonomy. Davenport et al. (2020) present a framework for comprehending AI's multidimensional impact on marketing strategies and customer behaviors, highlighting policy questions related to privacy, bias, and ethics and suggesting a role for AI augmentation. Other scholars focus on broader contexts. White and Boatwright (2020) discuss the ethical implications of using Facebook for public relations, considering the business model, algorithm-driven profiles, and deceptive communication strategies of the platform and assessing social responsibility implications. Gerlick and Liozu (2020) analyze legal and ethical dimensions of personalized pricing using algorithmic models, exploring intersections of antitrust, data privacy, and antidiscrimination frameworks with

considerations for deception, fairness, and social justice. Stanciu and Rîndaşu (2021) examine the practical implications of AI-based solutions in retail mobile applications, focusing on benefits, privacy concerns and required permissions.

Building an AI able to adapt to humans

This category involves almost 9% of the articles and deals with ethical concerns arising from AI's capacity to interact with humans in the workplace.

Effective human–AI interaction (6.6%). This research line addresses the ethical design of human–AI interactions. Miller (2019) contemplates the symbiotic relationship between humans and AI, discusses the impact of AI on various professions, and explores the concept of brain-computer interfaces. Gerdes (2018) highlights the need for inclusive ethical AI design, aligning AI with human values, and promoting moral growth in AI professionals. Another research line examines the frameworks needed to ensure an ethical human–AI interaction. Trunk et al. (2020) provide insights into integrating AI into organizational decision-making in situations of uncertainty. Like other researchers, they also emphasize the need for ethical frameworks within the context of education. Boni (2021) highlights the ethical dimension of human–AI collaboration, discussing the need for an adequate regulatory framework, human oversight, and AI digital literacy towards the ethical use of AI technologies.

Dialogue systems (2.3%). Under this section, scholars investigate user perceptions and expectations of AI in the workplace. Prakash and Das (2020) focus on user perceptions of AI-based conversational agents in mental healthcare services, analyzing factors influencing their adoption and use. Grimes et al. (2021) explore how users' expectations of conversational agents impact their evaluation, suggesting that user-formed expectations can influence perceptions beyond actual agent performance. Terblanche (2020) presents a design framework for creating AI coaches in organizational settings while adhering to coaching standards, ethics, and theoretical models. Tekin (2021) critically examines smartphone psychotherapy chatbots for mental illness diagnosis and treatment and discusses challenges related to early diagnosis, stigma, and global access to mental healthcare. Borau et al. (2021) investigate the perception of gendered chatbots, highlighting ethical questions regarding the humanization of AI based on gendered characteristics. Other scholars deal with societal implications of AI dialog systems. Mulvenna et al. (2021) explore ethical issues related to digital phenotyping, democratizing machine learning, and AI in digital health technologies. Berberich et al. (2020) propose incorporating the concept of harmony from East Asian cultures into the ethical discussion on AI, suggesting that by harmonizing AI, it will make intelligent systems tactful and sensitive to specific contexts.

Attributing the responsibility for AI's failures

This section, constituting almost 8% of the articles, addresses the implications arising from AI acting and learning without direct human supervision, encompassing two main issues: a responsibility gap and AI's moral status.

AI moral agency and legal status (5.1%). This research line consists of two main issues. The first one concerns the existence and status of artificial moral agency (AMAs). Nowik (2021) analyzes the legal and ethical implications of attributing electronic personhood to AI in employment relations by looking at concepts like AI as an employer, liability, and mandatory insurance. Kornai (2014) discusses the moral obligations of autonomous artificial general intelligences

(AGIs), as well as the challenges of bounding AGIs with ethical rationalism. Smith and Vickers (2021) examine how moral responsibility could be attributed to AI using a Strawsonian account.

Other researchers discuss the design of artificial moral agents. Mabaso (2021) discusses the use of exemplarism, an ethical theory, in building computationally rational AMAs. Gunkel (2014) advocates for including robots and AI in moral considerations and offers a critique of the limitations of current moral reasoning frameworks. Wallach (2010) stresses the need for a comprehensive model of moral decision-making in developing artificial moral agents, with a focus on mechanisms beyond traditional cognitive factors.

Responsibility gap (2.7%). This research reflects on the concept of the responsibility gap in AI, where an AI agent's actions that cause harm can lack clear responsibility. Saunders and Locke (2020) draw parallels between ancient practices of casting lots and AI in business decision-making and how, in both cases, control and moral responsibility are relinquished. Johnson (2015) discusses the potential emergence of a responsibility gap autonomous artificial agents of the future, emphasizing that responsibility allocation depends on human choices more than technological complexity. Awad et al. (2019) explore moral dilemmas in self-driving cars and propose that addressing these dilemmas requires collective discussions and agreements on ethical AI principles. Other scholars address responsibility gaps in AI systems, such as Santoni de Sio and Mecacci (2021), who identify interconnected responsibility gaps in AI and propose designing socio-technical systems for "meaningful human control" to comprehensively address these gaps. Schuelke-Leech et al. (2019) examine unexpected differences in the language used in policy documents and discussions about responsibility for highly automated vehicles.

Humans' unethical conducts

This category comprises over 2.5% of the articles and focuses on two key issues: the risk of exploiting ethics for economic gain and the peril of delegating tasks to AI that should inherently be human-centric.

Instrumental and perfunctory use of ethics (1.4%). This subset discusses the concern that principles and standards could be exploited for economic advantages, potentially leading companies to select countries and their markets where there are less stringent ethical regulations (Mikuriya et al., 2020). One research line endeavors to enhance the ethical impact in business. Rességuier and Rodrigues (2020) advocate for impactful ethical principles in AI and argue that relying on ethics as a replacement for legal frameworks poses the risk of its misapplication. Metcalf et al. (2019) explore the tension between industry commitments and operationalizing ethics within the tech sector and hold that ethics can either challenge or reinforce existing industry logics. Another line focuses on AI in Europe. Palladino (2021) explores how epistemic communities contribute to the constitutionalization of internet governance, analyzing the European Commission High-Level Expert Group on AI as a case study. He warns against the instrumental use of ethics in the ongoing debate, which he argues could lead a situation where self-governance is masked by ethical discourse. Bonsón et al. (2021) examine the inclusion of AI-related information and ethical principles in reports of European listed companies, focusing on AI system development, disclosure of ethical guidelines, and influencing factors.

Outsourcing human specificities (1.2%). Some papers within this subset deal with AI decision-making. In Danaher's (2018) examination of the ethics of using personal AI assistants, he

highlights concerns related to dehumanization while offering a nuanced view of the ethical implication of AI assistant use. Marie (2019) challenges the notion of human-algorithm complementarity in decision-making and raises concerns about algorithms influencing human decisions, particularly in domains like medicine. Ertemel et al. (2021) investigate the socio-economic consequences of AI, also raising concerns about outsourcing aspects of human life such as caregiving to machines, which could deprive society of the valuable dedication and spiritual benefits associated with human caregivers.

We now present the outcome of the preceding analysis through Figure 2, providing a qualitative overview of ethical concerns, with the most prevalent ones depicted at the top of the graphic and the less frequently occurring ones at the bottom.

The following section consolidates the solutions proposed by the literature to manage the ethical concerns raised by AI.

3.2 | Examining solutions and ethical approaches to AI ethical concerns

This section aims to map proposed solutions to AI ethical concerns, analyzing the specific ethical approaches that underpin those solutions. The classification is based on the distinction between the act- and agent-centered approaches described by Annas (1995). The rationale for adopting this classification is stems from several compelling reasons. First, it provides a clear method for categorizing the ethical approaches commonly used in AI ethics literature, whether they are implicit or explicit. Second, using well-established and recognized ethical categories promotes a more effective dialog between ethical theories—as conveyed through these specific ethical approaches—and the ever-evolving AI domain. Lastly, this framework further enhances the analysis by providing a structured examination of the solutions found in the literature. It sheds light on various facets through which ethical concerns tied to AI can be examined, thereby enriching the scope of the review.

Annas suggests that act-centered ethical theories primarily assess actions and decisions from an external observer point of view. Among these, two well-known ethical theories stand out. Utilitarianism, which evaluates the morality of an action based on the consequences: according to this ethical theory, the best solution is the one that achieves the greatest good for the greatest number of people. Second, deontological ethics determines the ethicality of an action based on principles or moral duties. In contrast, Annas also characterizes agent-centered ethical theories as those that place a stronger emphasis on the moral character of the agent. For instance, virtue ethics is a prominent agent-centered theory that focuses on the cultivation of virtuous character traits in individuals.

The two sections that follow categorize the solutions to the ethical concerns discussed in Section 3.1 as per this distinction, resulting in a classification of a solution as being underpinned by either an act-centered ethical approach (Section 3.2.1) or by agent-centered approach (Section 3.2.2).

3.2.1 | Act-centered ethics solutions

As explained above, consequentialism or utilitarianism evaluates how the consequences of actions maximize the good and minimize the harm, while deontological ethics centers on

FIGURE 2 The distribution of the ethical concerns

adhering to principles. Key indicators of these approaches include questions like, "which groups could be impacted by this choice?" or "how can these principles be applied here?" Such queries look at overall societal well-being and adopt an external perspective on the action.

There are three categories of solutions that underpin an act-centered ethical approach, namely, whether they follow guidelines and standards, whether they propose change(s) to the normative framework and whether they establish dedicated organizations.

Following guidelines and standards (43.5%)

The articles in this group adopt an act-centered approach, placing their primary focus on establishing and evaluating ethical procedures and guidelines for AI, such as frameworks, policies, and protocols, to ensure ethical behavior in AI-related practices. Furthermore, they propose new procedures, or adaptations to existing ones, with the aim of raising an awareness of ethics. Such an awareness would foster comprehensive policies to be developed that effectively address the ethical challenges posed by AI technologies and that can provide further guidance to AI practitioners and organizations.

A case in point is that of Thompson and Morgan (2020) who discuss AI's potential and its ethical challenges in the UK's National Health Service (NHS). They highlight the need for comprehensive policies, emphasizing actions and procedures in both AI development and its application. Siafakas (2021) suggests drawing inspiration from medical ethics to establish a Hippocratic Oath for AI scientists, which would cultivate a deeper ethical awareness among them.

Another line deals with impact assessment. These articles also extend their reach to assess AI's ethical implications across various domains, providing models and frameworks for an end-to-end evaluation. Sarkadi et al. (2019) addresses the challenge of deception in AI and proposes a model to analyze deceptive interactions. It focuses on assessing and influencing AI actions related to deception. Similarly, Wasilow and Thorpe (2019) introduce an ethics assessment framework for AI and robotics technologies in military applications.

All these proposals fall into the act-centered approach for two primary reasons. First, they emphasize shaping behavior through established models or norms rather than giving importance to the innate ethical qualities or personal development of AI scientists as individuals. Second, the core objective of proposing frameworks, policies, or oaths is to establish precise standards that oversee ethical conduct within the AI community. They focus on standardizing ethical behavior and actions rather than nurturing the ethical growth or intrinsic qualities of AI scientists themselves.

Change in the normative framework (19.4%)

Hin-Yan and Zawieska (2020) propose a human rights-based approach to complement responsible robotics initiatives, which strive to align human rights protections with the challenges posed by AI technologies. Their proposal looks at whether AI technologies ascribe to specific principles, more than looking at the acting agents from their own point of view. Fernandes et al. (2020) propose a theoretical model to understand conflicts arising from AI adoption and develop policies that benefit both adopters and non-adopters. The model considers policies and frameworks and focuses on AI's actions and their impact on various stakeholders. Hwang and Park (2020) assess how the Republic of Korea responds to AI threats by evaluating an AI Charter of Ethics, which examines whether AI-related actions and policies align with ethical standards. Fournier-Tombs (2021) advocates for a United Nations Regulation for AI, which

would set global ethical standards, while Meszaros and Ho (2021) explore how the EU General Data Protection Regulation (GDPR) applies to AI research.

The articles in this category are considered act-centered because they primarily focus on assessing the actions of AI systems and the entities responsible for their development and oversight. Specifically, their main emphasis lies in determining whether AI technologies and practices adhere to established ethical principles and normative frameworks.

Establishment of dedicated organizations (1.4%)

This group of articles proposes the establishment of dedicated organizations to manage the changes brought about by AI. This institutionalization is designed to govern and oversee the actions and behavior of AI within specific contexts. Leveringhaus (2018) advocates for an ethical framework for robot development. In particular, the author raises the concern that the potential risk of robots to infringe on human rights should be assessed and that appropriate institutions should be set up to hold humans accountable for the actions of robots. Similarly, Hamilton et al. (2021) propose the creation of an organization that utilizes a suggested rating system to educate users about how to choose a VAPA (Voice Activated Personal Assistants) to meet their needs. Such a system would convey the ethical criteria for choosing AI applications, highlighting the importance of evaluating AI actions and how they impact users' decisions.

To recap, these articles are act-centered for three reasons. First, they emphasize the creation of external bodies to regulate AI behavior. Second, they prioritize the scrutiny and evaluation of AI actions to ensure alignment with ethical principles and societal norms. Third, they involve the development of ethical standards and guidelines for AI systems and the monitoring of their adherence to these standards.

3.2.2 | Agent-centered ethics solutions

Agent-centered ethics takes the perspective of the individual, focusing on personal morality and the individual's character. These theories place emphasis on the cultivation of moral virtues and ethical excellence rather than merely assessing actions. They highlight the importance of being a virtuous person. A prominent example of an agent-centered ethical approach is virtue ethics, which adopts a first-person stance and asserts that *eudaimonia*, or flourishing, is the ultimate goal of human life (Aristotle, 1999). This school emphasizes personal improvement through consistent practice of virtues, fostering habitual alignment of mindset and actions. Unlike offering universal prescriptions, this approach relies on addressing queries such as "what actions contributes to my personal flourishing and the common good of society in this situation?" and "how can I be the best version of myself while dealing with AI-based technologies?" A virtue ethics approach to AI would involve placing a greater focus on the development of virtuous characteristics and behaviors in the design, development, and use of AI systems. These characteristics include empathy, compassion, fairness, and ethical judgment and would be aimed at AI developers, users, and stakeholders. The idea is to create AI systems that not only perform tasks efficiently but also align with ethical values and promote human well-being. Hence, virtue ethics encourages responsible AI development that considers the broader impact on society and aims to create AI-based systems that reflects and enhances virtuous qualities.

There are four categories of solutions that underpin an agent-centered ethical approach, namely, the involvement of stakeholders and communities, improving cooperation in AI research field, updating educational programs, and behaving virtuously.

Involve stakeholders and communities (18.6%)

These papers aim to involve stakeholders in AI processes and emphasize the role of individuals and communities (agents) in shaping ethical AI practices. Leikas et al. (2019) recognize the importance of engaging various agents in the decision-making process related to AI systems design, while Aizenberg and van den Hoven (2020) propose involving societal stakeholders in translating human rights into AI design. Du and Xie (2021) highlight the role of corporate social responsibility in shaping ethical AI in consumer markets. Aitken et al. (2020) focus on establishing a "Social License" for Financial Technology through public engagement. Since these articles explicitly emphasize the importance of engaging in dialog with stakeholders and involving the community, their specific focus lies in recognizing stakeholders and the community as essential participants in discussions about AI ethics. These papers offer different perspectives on how the ethical approval and acceptance of AI technologies depend on the collective decisions and perspectives of agents in the broader community. By making the responsibility of corporations and their interactions with consumers and other stakeholders their focal point, and how public engagement can shape AI design, these articles adopt an agent-centered ethical stance.

Improving cooperation in AI research field (8.3%)

The articles in this category also align with the agent-centered approach, since they are concerned with the actions and conduct of AI researchers. Some papers explore the transformations in research practices brought about by AI such as Goltz and Dondoli (2019), who argue that AI often takes precedence over other meaningful aspects of legal research. Their emphasis is on how AI tools affect research practices and ethical considerations, highlighting the roles and responsibilities of researchers in ensuring ethical conduct in AI research. Others introduce innovative research approaches, for example, Losbichler and Lehner (2021), who explore the complementary nature of human–machine information processing. Their article also presents a research agenda centered on human–machine collaboration and the impact and potential of AI in the realm of accounting leadership. This underscores the pivotal role of AI in research and its ethical implications. ÓhÉigeartaigh et al. (2020) addresses the challenges related to cross-cultural cooperation in AI ethics and governance. The authors underscore that misunderstandings between cultures and regions significantly hinder cross-cultural trust and argue that cooperation can be promoted and improved through both individual and collective actions. Furthermore, the article outlines practical measures and initiatives involving researchers and academia to foster cross-cultural cooperation, thereby highlighting the agency of individuals and institutions in advancing collaboration in AI ethics and governance. These contributions firmly fall within the agent-centered category as they center on the actions, behaviors, and cooperation of diverse cultural and regional agents, including researchers and policymakers.

Updating education programs (5.1%)

This group proposes updating education programs to address AI ethics and to bring the role and agency of individuals, educators, and learners in shaping ethical AI practices and education to the forefront. Nicola and Dalessio (2019) argue in favor of curricula that include the role of educators in integrating ethical considerations into the skill sets of future business professionals. Robinson and Bawden (2017) aim to integrate technical and ethical aspects into library/information education for data literacy. They acknowledge that educators play a pivotal role in equipping students with the ethical knowledge and skills necessary for responsible data practices. Liu and Murphy (2020) stress incorporating ethics and responsibility principles into

AI system design. They highlight the active role of AI system designers and developers in considering ethical dimensions especially during the design phase. Renz and Vladova (2021) advocate for integrating human-centered AI principles into educational technologies. They recognize the agency of educators and technology developers in developing educational AI systems that take ethical considerations into account. Dikici (2021) revises techniques for teaching ethics in engineering and assert that educators have the responsibility to adapt and improve pedagogical methods for teaching AI ethics effectively. Towards a similar aim, Burton et al. (2017) provide case studies for integrating ethics into AI courses. They emphasize the role of educators in structuring courses that actively engage students in ethical discussions and decision-making related to AI.

Behaving virtuously (3.7%)

This group includes articles that suggest the adoption of a virtue ethics approach to AI design and decision-making. Neubert and Montañez (2020) use Google as a case study to illustrate the potential of virtue ethics in steering ethical AI design and utilization. They highlight virtue as a foundational framework for ethical decision-making and its capacity to attract and retain personnel and customers while addressing ethical predicaments faced by organizations. Constantinescu et al. (2021) establish a link between the concept of responsible AI and Aristotelian virtue ethics. They furnish a conceptual model to address ethical quandaries in AI development, underscoring the role of dianoetic virtues and context in shaping moral accountability in AI systems. In a second line of research, the articles examine the ethical considerations regarding human–AI relationships and propose virtue ethics as a solution. Li (2021) evaluates the Friendly AI paradigm through the lens of virtue ethics, identifying four key issues. The study underscores the significance of human moral development. Kim et al. (2021) advocate for the infusion of virtue ethics to reinforce human control over AI systems. They propose a master–slave paradigm for AI, ensuring AI serves human interests instead of a scenario where AI dominates. Mabaso (2021) follows the exemplars approach, which is an ethical theory rooted in virtue ethics. It involves using moral exemplars, meeting community expectations, and practical simplicity to develop computationally rational Artificial Moral Agents (AMAs). Ratti and Graves (2021) argue that instead of relying solely on adherence to widely accepted principles, they advocate for cultivating moral virtues within the practice of data science. This perspective emphasizes that ethical decision-making involves the development of moral abilities or virtues through practice and posit the concept of "moral attention," which consists of the capability of data scientists to discern the ethical implications of their work.

Figure 3 synthetizes the corresponding act- and agent-centered solution for each ethical concern.

However, the literature focuses on some approaches more than others, as synthesized by Figure 4.

4 | DISCUSSION ON THE FINDINGS AND CONCLUSIONS

It is evident that analysis conducted here facilitates a deeper understanding of the current debate on the ethical concerns regarding AI and that some initial yet relevant conclusions can be drawn.

Responding to RQ1, the article provides a more exhaustive map of the ethical concerns than previous contributions. The measure of the frequency that ethical concerns are raised in the

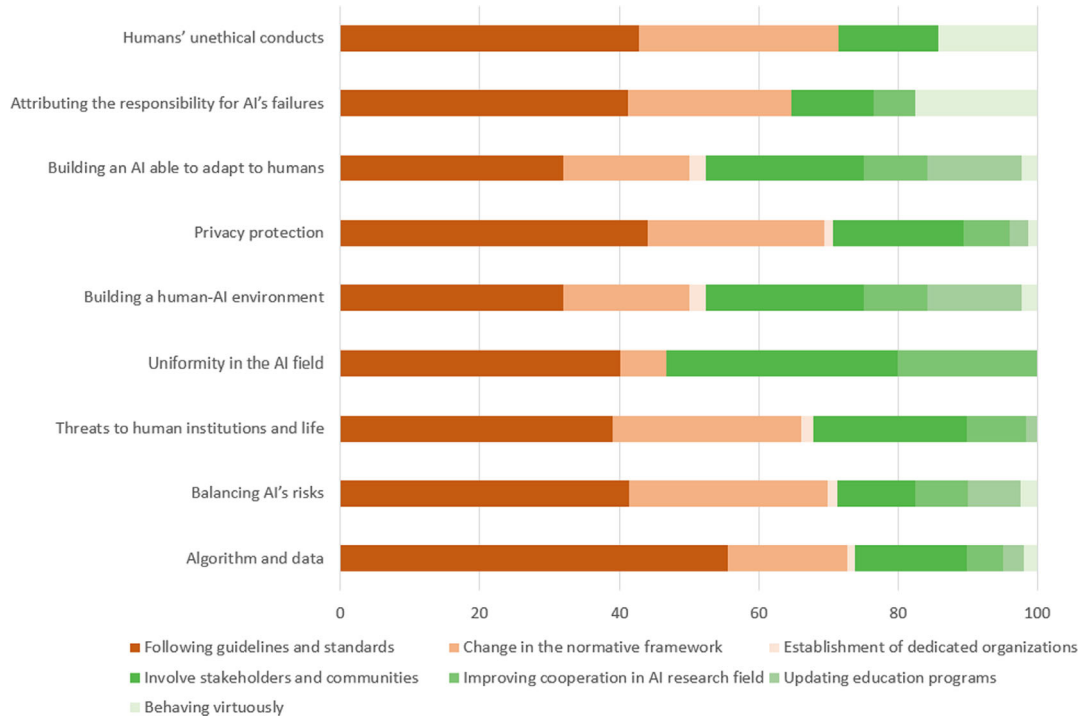

FIGURE 3 The distribution of act- and agent-centered proposed solutions among ethical concerns

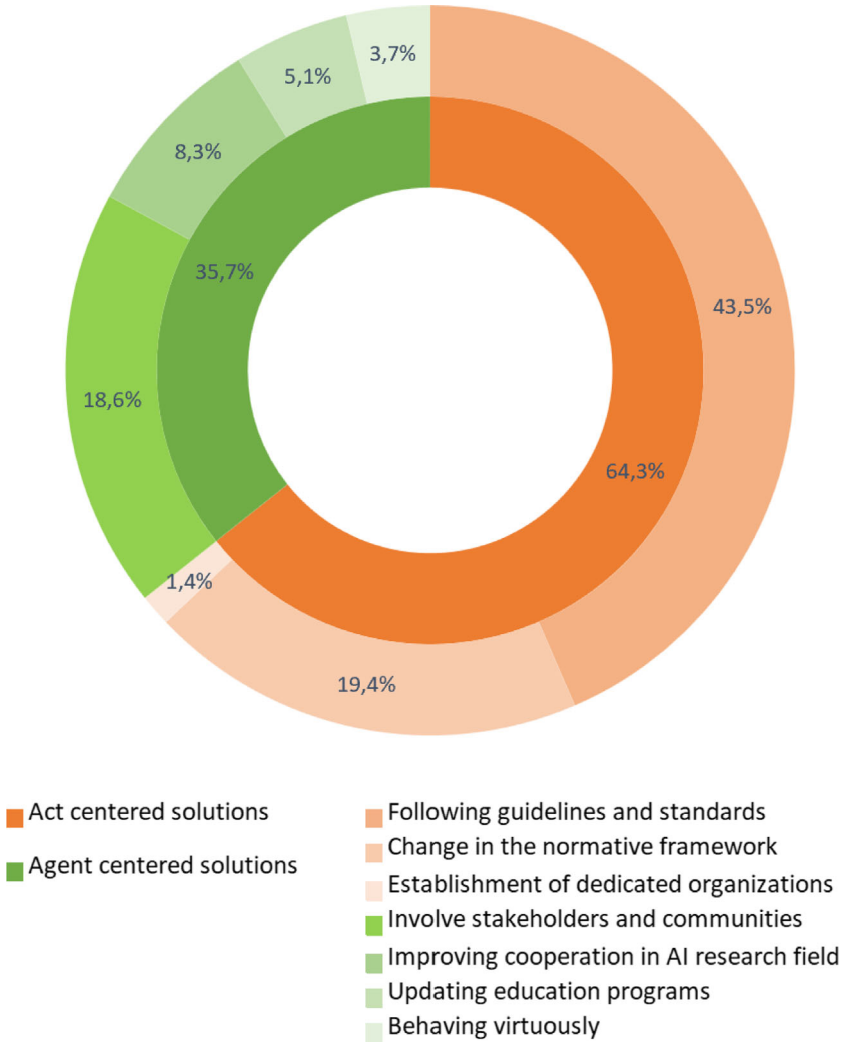

FIGURE 4 The distribution of the proposed solutions in governing AI's ethical concerns. [Correction added on 12 February 2024, after first online publication: Numerical values has been added on figure 4.]

literature offers valuable insights and points towards possible future research in two key areas. First, in analyzing the literature, making the prevalent ethical concerns visible aids in establishing research priorities that could channel efforts towards urgent and critical ethical aspects of AI. Concerning AI design, recurring ethical issues encompass algorithm transparency, data bias, algorithm fairness, design flaws, unpredictability, and threats to democratic values. Regarding human–AI interactions, identified concerns include business and job impacts, citizen risks, AI moral agency, and effective human–AI interaction. However, the research indicates that scarce attention has been paid to concerns that relate to "uniformity in the AI field" in design, "accessible AI," and the "instrumental and perfunctory use of ethics" in human–AI interactions. Yet these concerns remain relevant and warrant the attention of researchers because of their societal significance, particularly in terms of equality, fairness, and legal implications. Furthermore, in the future, this study could serve as a baseline to compare the literature

landscape against, which would tell us more about its evolution. Second, the research describes the relationship between the literature and the themes debated in society. While certain themes, such as AI takeover, draw significant attention in public discourse perhaps owing to the growing number of movies that are (mostly) about such dystopias, the literature appears less engrossed in this matter. Future research can shed light on the reasons why there might be a divergence (or convergence) of ethical concerns between academic discourse and societal debates.

Addressing RQ2 and RQ3, the study underscores a prevailing pattern observed in the solutions found in the existing literature. Generally, current literature aims to intervene by creating a shared space of rules and procedures for all actors involved in ethical issues concerning AI. Certainly, act-centered ethics approaches dominate agent-centered ethics in this regard. The solutions put forth to address ethical concerns are mostly based on an act-centered approach. Specifically, when addressing ethical concerns related to "algorithm and data," "humans' unethical conduct," "balancing AI's risks," and "privacy protection," the prevailing act-centered recommendation is to adhere to guidelines and standards. However, there is an exception concerning the "uniformity in the AI field" concern, where the proposed solutions tend to favor an agent-centered approach, particularly by involving stakeholders and communities in the process.

Still, an act-centered approach could present some criticalities. First, an eventual proliferation of interventions in the system's architecture could create a plethora of standards, models, and laws, which would overwhelm AI practitioners and companies. Besides, some authors already challenge the current set of solutions provided. Some assert that they could lack effectiveness and, hence, they explicitly aim to complement the regulation and link theory to practice (Aitken et al., 2020), and Grodzinsky (2017), for example, argues that some formulas work in specific environments, but not in others. Besides studying the impact of ethical guidelines on ethical decision-making for software engineers, McNamara et al. (2018) question the effectiveness of the instruments in changing the behavior, and finally with reference to the field of HR.

After this analysis, it is evident that the agent-centered solutions, and among them virtue ethics, have not attracted considerable attention within the field so far. Nonetheless, some valuable considerations arise. First, advocating for this approach does not mean disregarding other emerging viewpoints from the existing literature. Alasdair MacIntyre, a prominent figure in virtue ethics (Ferrero & Sison, 2014), has integrated norms into his reflections, complementing them with goods and virtues (MacIntyre, 1992). Even though virtue ethics is agent-centered and focuses on the moral character of the agent as well as the importance of cultivating virtues that serve as a moral compass for individuals when facing ethical dilemmas, it also integrates norms and consequences. Virtue ethics acknowledges that ethical norms and principles are essential guides for action, but it places them within the broader context of character development. Moreover, it can help AI developers and users identify and address potential ethical blind spots that may not be evident through a rule-based approach. While virtue ethics places character at the forefront, it does not ignore consequences. Since its ultimate goal is to achieve human flourishing or *eudaimonia*, virtue ethics promotes that agents contemplate the implications of their actions on their own character, the character of others, and the overall well-being of society. Therefore, it encourages AI developers and users to think about the long-term consequences of AI systems on humanity.

Second, the virtue ethics framework has the capacity to encompass elements such as character development and emotions. Considering the complexity of human nature holds multifaceted benefits. As AI increasingly interacts with humans, the incorporation of emotions and human subjectivity in the analysis can lead to AI models that mirror real interactions, hence contributing to the acceptance of AI in society. Virtue ethics emphasizes the cultivation of moral virtues

in individuals involved in AI development and use. This can lead to a more ethically aware and responsible AI community. Moreover, adopting this perspective would enhance discussions around pressing matters and the concerns spoken about in this article. As an illustration, consider the ongoing discourse about artificial moral agents (AMAs). Here, the framework of virtue ethics has the potential to enhance our understanding significantly, thanks to its all-encompassing perspective on "interactivity, autonomy, adaptation, consciousness, intentionality, free will" (Sison & Redín, 2021, p. 16). In particular, backed by a fully articulated philosophical anthropology, the virtue ethics approach provides a theoretical framework that offers clarity, depth, and coherence to the discussion (Sison & Redín, 2021). In addition, virtue ethics does not rely on rigid rules or algorithms, allowing for a more adaptable approach to addressing complex ethical dilemmas in AI. It thereby encourages nuanced ethical decision-making. Since virtue ethics places a strong emphasis on human values and well-being, it helps to ensure that AI is designed and used in ways that prioritize human priorities and dignity, at the individual and social level. In sum, virtue ethics offers a holistic, adaptable, and character-focused framework for addressing ethical concerns in AI, which promotes responsible and ethical AI development and use. Meanwhile, the field is already dynamic, fostering discussions like the potential for virtuous machines, championed by proponents (Gamez et al., 2020) and critiqued by detractors (Constantinescu & Crisp, 2021). Future studies could delve into crafting practical, virtue ethics-centered solutions that tackle specific ethical concerns, further enriching the discourse.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare that is relevant to the content of this article.

ETHICS STATEMENT

This research does not involve human participants or animals.

ORCID

Francesco Vincenzo Giarmoleo https://orcid.org/0009-0005-6139-0481
Ignacio Ferrero https://orcid.org/0000-0001-5869-8366
Marta Rocchi https://orcid.org/0000-0003-4668-7445
Massimiliano Matteo Pellegrini https://orcid.org/0000-0002-5722-2988

ENDNOTE

[1] The OECD (2019, p. 7) gives the following definition of AI: "An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy."

REFERENCES

Aitken, M., Toreini, E., Carmichael, P., Coopamootoo, K., Elliott, K., & van Moorsel, A. (2020). Establishing a social licence for financial technology: Reflections on the role of the private sector in pursuing ethical data practices. *Big Data & Society*, *7*(1), 2053951720908892. https://doi.org/10.1177/2053951720908892

Aizenberg, E., & van den Hoven, J. (2020). Designing for human rights in AI. *Big Data & Society*, *7*(2), 205395172094956. https://doi.org/10.1177/2053951720949566

Alderman, J. (2021). Women in the smart machine age: Addressing emerging risks of an increased gender gap in the accounting profession. *Journal of Accounting Education*, *55*, 100715. https://doi.org/10.1016/j.jaccedu.2021.100715

Annas, J. (1995). *The morality of happiness (first paperback edition)*. Oxford University Press. https://doi.org/10.1093/0195096525.001.0001

Aristotle. (1999). *Nicomachean ethics*. Hackett Publishing Co.

Arnold, T., & Scheutz, M. (2018). The "big red button" is too late: An alternative model for the ethical evaluation of AI systems. *Ethics and Information Technology*, *20*, 59–69. https://doi.org/10.1007/s10676-018-9447-7

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management*, *62*, 102433. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Astobiza, A. M., Toboso, M., Aparicio, M., & Lopez, D. (2021). AI ethics for sustainable development goals. *IEEE Technology & Society Magazine*, *40*(2), 66–71. https://doi.org/10.1109/MTS.2021.3056294

Awad, E., Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2019). The thorny challenge of making moral machines: Ethical dilemmas with self-driving cars. *NIM Marketing Intelligence Review*, *11*(2), 42–47. https://doi.org/10.2478/nimmir-2019-0015

Bao, J., & Xie, Q. (2022). Artificial intelligence in animal farming: A systematic literature review. *Journal of Cleaner Production*, *331*, 129956. https://doi.org/10.1016/j.jclepro.2021.129956

Belk, R. (2021). Ethical issues in service robotics and artificial intelligence. *Service Industries Journal*, *41*(13/14), 860–876. https://doi.org/10.1080/02642069.2020.1727892

Benanti, P. (2019). Le macchine sapienti: Intelligenze artificiali e decisioni umane. Marietti 1820.

Berberich, N., Nishida, T., & Suzuki, S. (2020). Harmonizing artificial intelligence for social good. *Philosophy & Technology*, *33*(4), 613–638. https://doi.org/10.1007/s13347-020-00421-8

Boni, M. (2021). The ethical dimension of human–artificial intelligence collaboration. *European View*, *20*(2), 182–190. https://doi.org/10.1177/17816858211059249

Bonsón, E., Lavorato, D., Lamboglia, R., & Mancini, D. (2021). Artificial intelligence activities and ethical approaches in leading listed companies in the European Union. *International Journal of Accounting Information Systems*, *43*, 100535. https://doi.org/10.1016/j.accinf.2021.100535

Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most human bot: Female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing*, *38*(7), 1052–1068. https://doi.org/10.1002/mar.21480

Borges, A. F., Laurindo, F. J., Spínola, M. M., Gonçalves, R. F., & Mattos, C. A. (2021). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *International Journal of Information Management*, *57*, 102225. https://doi.org/10.1016/j.ijinfomgt.2020.102225

Brown, S., Davidovic, J., & Hasan, A. (2021). The algorithm audit: Scoring the algorithms that score us. *Big Data & Society*, *8*(1), 205395172098386. https://doi.org/10.1177/2053951720983865

Burr, C., Taddeo, M., & Floridi, L. (2020). The ethics of digital well-being: A thematic review. *Science and Engineering Ethics*, *26*, 2313–2343. https://doi.org/10.1007/s11948-020-00175-8

Burton, E., Goldsmith, J., Koenig, S., Kuipers, B., Mattei, N., & Walsh, T. (2017). Ethical considerations in artificial intelligence courses. *AI Magazine*, *38*(2), 22–34. https://doi.org/10.1609/aimag.v38i2.2731

Busalim, A. H. (2016). Understanding social commerce: A systematic literature review and directions for further research. *International Journal of Information Management*, *36*(6), 1075–1088. https://doi.org/10.1016/j.ijinfomgt.2016.06.005

Callanan, G. A., Perri, D. F., & Tomkowicz, S. M. (2021). Targeting vulnerable populations: The ethical implications of data mining, automated prediction, and focused marketing. *Business and Society Review*, *126*(2), 155–167. https://doi.org/10.1111/basr.12233

Carman, M., & Rosman, B. (2021). Applying a principle of explicability to AI research in Africa: Should we do it? *Ethics and Information Technology*, *23*(2), 107–117. https://doi.org/10.1007/s10676-020-09534-2

Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., da Francisco, R. P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, *137*, 106024. https://doi.org/10.1016/j.cie.2019.106024

Cavallone, M., & Palumbo, R. (2020). Debunking the myth of industry 4.0 in health care: Insights from a systematic literature review. *The TQM Journal*, *32*(4), 849–868. https://doi.org/10.1108/TQM-10-2019-0245

Chahal, D., & Byrne, M. F. (2020). A primer on artificial intelligence and its application to endoscopy. *Gastrointestinal Endoscopy*, *92*(4), 813–820. https://doi.org/10.1016/j.gie.2020.04.074

Chen, Y. Y., Li, Y., & Li, C. J. (2020, September). Could artificial intelligence make human doctors obsolete? A survey based on public attitudes. *European Journal of Public Health*, *30*(Supplement_5), ckaa166.040. https://doi.org/10.1093/eurpub/ckaa166.040

Chiao, V. (2019). Fairness, accountability and transparency: Notes on algorithmic decision-making in criminal justice. *International Journal of Law in Context*, *15*(2), 126–139. https://doi.org/10.1017/S1744552319000077

Coeckelbergh, M. (2021). Time machines: Artificial intelligence, process, and narrative. *Philosophy & Technology*, *34*(4), 1623–1638. https://doi.org/10.1007/s13347-021-00479-y

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, *60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Constantinescu, M., & Crisp, R. (2021). Can robotic AI systems be virtuous and why does this matter? https://www.researchgate.net/publication/344072143_Can_robotic_AI_systems_be_virtuous_and_why_does_this_matter

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*(4), 803–814. https://doi.org/10.1007/s10676-021-09616-9

Danaher, J. (2018). Toward an ethics of AI assistants: An initial framework. *Philosophy & Technology*, *31*(4), 629–653. https://doi.org/10.1007/s13347-018-0317-3

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

de Almeida, P. G. R., dos Santos, C. D., & Farias, J. S. (2021). Artificial intelligence regulation: A framework for governance. *Ethics and Information Technology*, *23*(3), 505–525. https://doi.org/10.1007/s10676-021-09593-z

di Vaio, A., Palladino, R., Hassan, R., & Escobar, O. (2020). Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research*, *121*, 283–314. https://doi.org/10.1016/j.jbusres.2020.08.019

Dignum, V. (2018). Ethics in artificial intelligence: Introduction to the special issue. *Ethics and Information Technology*, *20*(1), 1–3. https://doi.org/10.1007/s10676-018-9450-z

Dikici, B. (2021). Implementing next-generation engineering ethics in education and industry for social happiness: A macroscopic overview from Turkey to the world. *Turkish Journal of Business Ethics*, *14*(2), 364–380. https://doi.org/10.12711/tjbe.2021.14.2.2303

Donia, J., & Shaw, J. A. (2021). Co-design and ethical artificial intelligence for health: An agenda for critical research and practice. *Big Data & Society*, *8*(2), 205395172110652. https://doi.org/10.1177/20539517211065248

Du, S., & Xie, C. (2021). Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *Journal of Business Research*, *129*, 961–974. https://doi.org/10.1016/j.jbusres.2020.08.024

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., Jain, V., Karjaluoto, H., Kefi, H., Krishen, A. S., Kumar, V., Rahman, M. M., Raman, R., Rauschnabel, P. A., Rowley, J., Salo, J., Tran, G. A., & Wang, Y. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, *59*, 102168. https://doi.org/10.1016/j.ijinfomgt.2020.102168

Ertemel, A. V., Karadayı, T., & Makaritou, P. (2021). Investigating the socio-economic consequences of artificial intelligence: A qualitative research. *Journal of International Trade, Logistics and Law*, *7*(1), 75–89.

European Parliament. (2023). DRAFT Compromise Amendments on the Draft Report. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/CJ40/DV/2023/05-11/ConsolidatedCA_IMCOLIBE_AI_ACT_EN.pdf

Fernandes, P. M., Santos, F. C., & Lopes, M. (2020). Norms for beneficial a.I.: A computational analysis of the societal value alignment problem. *AI Communications*, *33*(3–6), 155–171. https://doi.org/10.3233/AIC-201502

Ferrero, I., & Sison, A. J. G. (2014). A quantitative analysis of authors, schools and themes in virtue ethics articles in business ethics and Management Journals (1980-2011). *Business Ethics: A European Review*, *23*(4), 375–400. https://doi.org/10.1111/beer.12057

Formosa, P., Wilson, M., & Richards, D. (2021). A principlist framework for cybersecurity ethics. *Computers & Security*, *109*, 102382. https://doi.org/10.1016/j.cose.2021.102382

Fournier-Tombs, E. (2021). Towards a United Nations internal regulation for artificial intelligence. *Big Data & Society*, *8*(2), 205395172110394. https://doi.org/10.1177/20539517211039493

Fritts, M., & Cabrera, F. (2021). AI recruitment algorithms and the dehumanization problem. *Ethics and Information Technology*, *23*(4), 791–801. https://doi.org/10.1007/s10676-021-09615-w

Gabriel, I. (2020). Artificial intelligence, values, and alignment. *Minds and Machines*, *30*(3), 411–437. https://doi.org/10.1007/s11023-020-09539-2

Gamez, P., Shank, D. B., Arnold, C., & North, M. (2020). Artificial virtue: The machine question and perceptions of moral character in artificial moral agents. *Ai & Society*, *35*(4), 795–809. https://doi.org/10.1007/s00146-020-00977-1

Gerdes, A. (2018). An inclusive ethical design perspective for a flourishing future with artificial intelligent systems. *European Journal of Risk Regulation*, *9*(4), 677–689. https://doi.org/10.1017/err.2018.62

Gerlick, J. A., & Liozu, S. M. (2020). Ethical and legal considerations of artificial intelligence and algorithmic decision-making in personalized pricing. *Journal of Revenue and Pricing Management*, *19*(2), 85–98. https://doi.org/10.1057/s41272-019-00225-2

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Goltz, N., & Dondoli, G. (2019). A note on science, legal research and artificial intelligence. *Information & Communications Technology Law*, *28*(3), 239–251. https://doi.org/10.1080/13600834.2019.1644065

Gomes, P., Verçosa, L., Melo, F., Silva, V., Filho, C. B., & Bezerra, B. (2022). Artificial intelligence-based methods for business processes: A systematic literature review. *Applied Sciences*, *12*(5), 2314. https://doi.org/10.3390/app12052314

Guo, J., Farhang-Razi, V., & Algra, P. R. (2019). AI: A glossary of terms. In E. R. Ranschaert, S. Morozov, & P. R. Algra (Eds.), *Artificial intelligence in medical imaging: opportunities, applications, and risks* (pp. 347–373). Springer.

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. Proceedings of the 52nd Hawaii international conference on system sciences.

Grimes, G. M., Schuetzler, R. M., & Giboney, J. S. (2021). Mental models and expectation violations in conversational AI interactions. *Decision Support Systems*, *144*, 113515. https://doi.org/10.1016/j.dss.2021.113515

Grodzinsky, F. S. (2017). Why big data needs the virtues. In *Philosophy and computing* (pp. 221–234). Springer. https://doi.org/10.1007/978-3-319-61043-6_12

Guha, A., Grewal, D., Kopalle, P. K., Haenlein, M., Schneider, M. J., Jung, H., Moustafa, R., Hegde, D. R., & Hawkins, G. (2021). How artificial intelligence will affect the future of retailing. *Journal of Retailing*, *97*(1), 28–41. https://doi.org/10.1016/j.jretai.2021.01.005

Gunkel, D. J. (2014). A vindication of the rights of machines. *Philosophy & Technology*, *27*(1), 113–132. https://doi.org/10.1007/s13347-013-0121-z

Hagerty, A., & Rubinov, I. (2019). Global AI ethics: A review of the social impacts and ethical implications of artificial intelligence. arXiv preprint arXiv:1907.07892.

Hamilton, C., Swart, W., & Stokes, G. M. (2021). Developing a measure of social, ethical, and legal content for intelligent cognitive assistants. *Journal of Strategic Innovation and Sustainability*, *16*(3). https://doi.org/10.33423/jsis.v16i3.4438

Harambam, J., Helberger, N., & van Hoboken, J. (2018). Democratizing algorithmic news recommenders: How to materialize voice in a technologically saturated media ecosystem. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences*, *376*(2133), 20180088. https://doi.org/10.1098/rsta.2018.0088

Helberger, N., Araujo, T., & de Vreese, C. H. (2020). Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making. *Computer law & Security Review*, *39*, 105456. https://doi.org/10.1016/j.clsr.2020.105456

Hickman, E., & Petrin, M. (2021). Trustworthy AI and corporate governance: The EU's ethics guidelines for trustworthy artificial intelligence from a company law perspective. *European Business Organization Law Review*, *22*(4), 593–625. ABI/INFORM Collection. https://doi.org/10.1007/s40804-021-00224-0

Hin-Yan, L., & Zawieska, K. (2020). From responsible robotics towards a human rights regime oriented to the challenges of robotics and artificial intelligence. *Ethics and Information Technology*, *22*(4), 321–333. https://doi.org/10.1007/s10676-017-9443-3

Hongladarom, S. (2021). The Thailand national AI ethics guideline: An analysis. *Journal of Information, Communication & Ethics in Society*, *19*(4), 480–491. ABI/INFORM Collection. https://doi.org/10.1108/JICES-01-2021-0005

Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, *178*, 977–1007. https://doi.org/10.1007/s10551-022-05049-6

Hwang, H., & Park, M. H. (2020). The threat of AI and our response: The AI charter of ethics in South Korea. *Asian Journal of Innovation & Policy*, *9*(1), 56–78.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly*, *37*(3), 101493. https://doi.org/10.1016/j.giq.2020.101493

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Johnson, D. G. (2015). Technology with no human responsibility? *Journal of Business Ethics*, *127*(4), 707–715. https://doi.org/10.1007/s10551-014-2180-1

Kamishima, Y., Gremmen, B., & Akizawa, H. (2018). Can merging a capability approach with effectual processes help us define a permissible action range for AI robotics entrepreneurship? *Philosophy of Management*, *17*(1), 97–113. Philosopher's Index

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 Preprint ArXiv:2109.07906.

Kim, T. W., Maimone, F., Pattit, K., Sison, A. J., & Teehankee, B. (2021). Master and slave: The dialectic of human-artificial intelligence engagement. *Humanistic Management Journal*, *6*(3), 355–371. https://doi.org/10.1007/s41463-021-00118-w

Kornai, A. (2014). Bounding the impact of AGI. *Journal of Experimental & Theoretical Artificial Intelligence*, *26*(3), 417–438. https://doi.org/10.1080/0952813X.2014.895109

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology*. SAGE Publications, Inc.. https://doi.org/10.4135/9781071878781

Leikas, J., Koivisto, R., & Gotcheva, N. (2019). Ethical framework for designing autonomous intelligent systems. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(1), 18. https://doi.org/10.3390/joitmc5010018

Leveringhaus, A. (2018). Developing robots: The need for an ethical framework. *European View*, *17*(1), 37–43. https://doi.org/10.1177/1781685818761016

Li, O. (2021). Problems with "friendly AI". *Ethics and Information Technology*, *23*(3), 543–550. https://doi.org/10.1007/s10676-021-09595-x

Lilkov, D. (2021). Regulating artificial intelligence in the EU: A risky game. *European View*, *20*(2), 166–174. https://doi.org/10.1177/17816858211059248

Liu, X. (Michelle), & Murphy, D. (2020). A multi-faceted approach for trustworthy AI in cybersecurity. *Journal of Strategic Innovation and Sustainability*, *15*(6). https://doi.org/10.33423/jsis.v15i6.3596

Losbichler, H., & Lehner, O. M. (2021). Limits of artificial intelligence in controlling and the ways forward: A call for future accounting research. *Journal of Applied Accounting Research*, *22*(2), 365–382. https://doi.org/10.1108/JAAR-10-2020-0207

Mabaso, B. A. (2021). Artificial moral agents within an ethos of AI4SG. *Philosophy & Technology*, *34*, 7–21. https://doi.org/10.1007/s13347-020-00400-z

MacIntyre, A. (1992). Plain persons and moral philosophy: Rules, virtues and goods. *American Catholic Philosophical Quarterly*, *66*(1), 3–19. https://doi.org/10.5840/acpq199266144

Mantelero, A., & Esposito, M. S. (2021). An evidence-based methodology for human rights impact assessment (HRIA) in the development of AI data-intensive systems. *Computer law & Security Review*, *41*, 105561. https://doi.org/10.1016/j.clsr.2021.105561

Marie, D. (2019). AI and the illusion of human-algorithm complementarity. *Social Research*, *86*(4), 887–908. https://doi.org/10.1353/sor.2019.0051

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*, 835–850. https://doi.org/10.1007/s10551-018-3921-3

Mathew, T., & Mathew, A. (2021). Ethical dilemma in future warfare—Use of automated weapon systems. *AIMS International Journal of Management*, *15*(3), 169–179. https://doi.org/10.26573/2021.15.3.3

McLean, S., Read, G. J. M., Thompson, J., Baber, C., Stanton, N. A., & Salmon, P. M. (2021). The risks associated with artificial general intelligence: A systematic review. *Journal of Experimental & Theoretical Artificial Intelligence*, *1–15*, 649–663. https://doi.org/10.1080/0952813X.2021.1964003

McNamara, A., Smith, J., & Murphy-Hill, E. (2018). Does ACM's code of ethics change ethical decision making in software development? 729–733.

McStay, A. (2020). Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy. *Big Data & Society*, *7*(1), 205395172090438. https://doi.org/10.1177/2053951720904386

Meek, T., Barham, H., Beltaif, N., Kaadoor, A., & Akhter, T. (2016). Managing the ethical and risk implications of rapid advances in artificial intelligence: A literature review. In *Portland international conference on management of engineering and technology (PICMET)* (pp. 682–693). IEEE.

Melé, D. (2021). Ethics at the workplace in the fourth industrial revolution: A Catholic social teaching perspective. *Business Ethics, the Environment & Responsibility*, *30*(4), 772–783. https://doi.org/10.1111/beer.12368

Meszaros, J., & Ho, C. (2021). AI research and data protection: Can the same rules apply for commercial and academic research under the GDPR? *Computer law & Security Review*, *41*, 105532. https://doi.org/10.1016/j.clsr.2021.105532

Metcalf, J., Moss, E., & Boyd, D. (2019). Owning ethics: Corporate logics, silicon valley, and the institutionalization of ethics. *Social Research*, *86*(2), 449–476. https://doi.org/10.1353/sor.2019.0022

Mikuriya, K., Cantens, T., Wolffgang, H.-M., Rogmann, A., Harden, K., Chapa, R. T., Widdowson, D., Blegen, B. C., Hesketh, D., & Peteva, J. (2020). If algorithms dream of customs, do customs officials dream of algorithms? A manifesto for data mobilisation in customs. *World Customs Journal*, *14*(2), 3–22.

Miller, A. (2019). The intrinsically linked future for human and artificial intelligence interaction. *Journal of Big Data*, *6*(1), 1–9. https://doi.org/10.1186/s40537-019-0202-7

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., & Berlin, J. A. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement (Chinese edition). *Journal of Chinese Integrative Medicine*, *7*(9), 889–896. https://doi.org/10.3736/jcim20090918

Möllmann, N. R., Mirbabaie, M., & Stieglitz, S. (2021). Is it alright to use artificial intelligence in digital health? A systematic literature review on ethical considerations. *Health Informatics Journal*, *27*(4). https://doi.org/10.1177/14604582211052391

Montes, G. A., & Goertzel, B. (2019). Distributed, decentralized, and democratized artificial intelligence. *Technological Forecasting and Social Change*, *141*, 354–358. https://doi.org/10.1016/j.techfore.2018.11.010

Morley, J., Machado, C. C., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social Science & Medicine*, *260*, 113172. https://doi.org/10.1016/j.socscimed.2020.113172

Mulvenna, M. D., Bond, R., Delaney, J., Dawoodbhoy, F. M., Boger, J., Potts, C., & Turkington, R. (2021). Ethical issues in democratizing digital phenotypes and machine learning in the next generation of digital health technologies. *Philosophy & Technology*, *34*(4), 1945–1960. https://doi.org/10.1007/s13347-021-00445-8

Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. (2020). The ethical implications of using artificial intelligence in auditing. *Journal of Business Ethics*, *167*(2), 209–234. https://doi.org/10.1007/s10551-019-04407-1

Mustak, M., Salminen, J., Plé, L., & Wirtz, J. (2021). Artificial intelligence in marketing: Topic modeling, scientometric analysis, and research agenda. *Journal of Business Research*, *124*, 389–404. https://doi.org/10.1016/j.jbusres.2020.10.044

Mutasa, S., Sun, S., & Ha, R. (2020). Understanding artificial intelligence based radiology studies: What is overfitting? *Clinical Imaging*, *65*, 96–99. https://doi.org/10.1016/j.clinimag.2020.04.025

Neubert, M. J., & Montañez, G. D. (2020). Virtue as a framework for the design and use of artificial intelligence. *Business Horizons*, *63*(2), 195–204. https://doi.org/10.1016/j.bushor.2019.11.001

Nicola, C. B., & Dalessio, D. (2019). Artificial intelligence and the impact on business curricula. *Academy of Business Research Journal*, *3*, 30–53.

Nowik, P. (2021). Electronic personhood for artificial intelligence in the workplace. *Computer law & Security Review*, *42*, 105584. https://doi.org/10.1016/j.clsr.2021.105584

OECD. (2019). Recommendation of the council on artificial intelligence. *OECD Legal Instruments*. https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449#:~:text=The%20OECD%20Recommendation%20on%20AI,governments%20in%20their%20implementation%20efforts

ÓhÉigeartaigh, S. S., Whittlestone, J., Liu, Y., Zeng, Y., & Liu, Z. (2020). Overcoming barriers to cross-cultural cooperation in AI ethics and governance. *Philosophy & Technology*, *33*(4), 571–593. https://doi.org/10.1007/s13347-020-00402-x

Oleksy, W., Just, E., & Zapedowska-Kling, K. (2012). Gender issues in information and communication technologies (ICTs). *Journal of Information, Communication & Ethics in Society*, *10*(2), 107–120. https://doi.org/10.1108/14779961211227010

Orr, W., & Davis, J. L. (2020). Attributions of ethical responsibility by artificial intelligence practitioners. *Information, Communication & Society*, *23*(5), 719–735. https://doi.org/10.1080/1369118X.2020.1713842

Pak-Hang, W. (2020). Cultural differences as excuses? Human rights and cultural values in global ethics and governance of AI. *Philosophy & Technology*, *33*(4), 705–715. https://doi.org/10.1007/s13347-020-00413-8

Palladino, N. (2021). The role of epistemic communities in the "constitutionalization" of internet governance: The example of the European Commission High-Level Expert Group on Artificial Intelligence. *Telecommunications Policy*, *45*(6), 102149. https://doi.org/10.1016/j.telpol.2021.102149

Pandya, J. (2019). Can artificial intelligence be biased? Forbes. Retrieved May 19, 2019, from https://www.forbes.com/sites/cognitiveworld/2019/01/20/can-artificial-intelligence-bebiased/#6c5538467e7c

Prakash, A. V., & Das, S. (2020). Intelligent conversational agents in mental healthcare services: A thematic analysis of user perceptions. *Pacific Asia Journal of the Association for Information Systems*, *12*(2), 1. https://doi.org/10.17705/1pais.12201

Quinn, R. A. (2021). Artificial intelligence and the role of ethics. *Statistical Journal of the IAOS*, Business Source Ultimate, *37*, 75–77. https://doi.org/10.3233/SJI-210791

Rakowski, R., Polak, P., & Kowalikova, P. (2021). Ethical aspects of the impact of AI: The status of humans in the era of artificial intelligence. *Society*, *58*(3), 196–203. https://doi.org/10.1007/s12115-021-00586-8

Ratti, E., & Graves, M. (2021). Cultivating moral attention: A virtue-oriented approach to responsible data science in healthcare. *Philosophy & Technology*, *34*(4), 1819–1846. https://doi.org/10.1007/s13347-021-00490-3

Reijers, W., Wright, D., Brey, P., Weber, K., Rodrigues, R., O'Sullivan, D., & Gordijn, B. (2018). Methods for practising ethics in research and innovation: A literature review, critical analysis and recommendations. *Science and engineering ethics*, *24*, 1437–1481.

Reim, W., Åström, J., & Eriksson, O. (2020). Implementation of artificial intelligence (AI): A roadmap for business model innovation. *AI*, *1*(2), 11–191. https://doi.org/10.3390/ai1020011

Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. *European Journal of Operational Research*, *291*(3), 906–917. https://doi.org/10.1016/j.ejor.2020.09.020

Renz, A., & Vladova, G. (2021). Reinvigorating the discourse on human-centered artificial intelligence in educational technologies. *Technology Innovation Management Review*, *11*(5), 5–16. https://doi.org/10.22215/timreview/1438

Rességuier, A., & Rodrigues, R. (2020). AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data & Society*, *7*(2), 205395172094254. https://doi.org/10.1177/2053951720942541

Robinson, L., & Bawden, D. (2017). 'The story of data': A socio-technical approach to education for the data librarian role in the CityLIS library school at City, University of London. *Library Management*, *38*(6/7), 312–322. Library & Information Science Collection. https://doi.org/10.1108/LM-01-2017-0009

Rosenberger, R., & Verbeek, P. P. (2015). *Postphenomenological investigations: Essays on human-technology relations*. Lexington Books.

Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News*, *27*, 31–39. https://doi.org/10.1108/01409170410784185

Sako, M. (2020). Technology strategy and management: Artificial intelligence and the future of professional work. *Association for Computing Machinery. Communications of the ACM*, *63*(4), 25–27. https://doi.org/10.1145/3382743

Santoni de Sio, F., & Mecacci, G. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology*, *34*(4), 1057–1084. https://doi.org/10.1007/s13347-021-00450-x

Sari, O., & Celik, S. (2021). Legal evaluation of the attacks caused by artificial intelligence-based lethal weapon systems within the context of Rome statute. *Computer Law & Security Review*, *42*, 105564. https://doi.org/10.1016/j.clsr.2021.105564

Sarkadi, Ş., Panisson, A. R., Bordini, R. H., McBurney, P., Parsons, S., & Chapman, M. (2019). Modelling deception using theory of mind in multi-agent systems. *AI Communications*, *32*(4), 287–302. https://doi.org/10.3233/AIC-190615

Saunders, K. T., & Locke, L. G. (2020). Casting lots, gambling, and artificial intelligence. *Journal of Biblical Integration in Business*, *23*(1), 58–69.

Schiff, D., Rakova, B., Ayesh, A., Fanti, A., & Lennon, M. (2021). Explaining the principles to practices gap in AI. *IEEE Technology & Society Magazine*, *40*(2), 81–94. https://doi.org/10.1109/MTS.2021.3056286

Schuelke-Leech, B., Jordan, S. R., & Barry, B. (2019). Regulating autonomy: An assessment of policy language for highly automated vehicles. *Review of Policy Research*, *36*(4), 547–579. https://doi.org/10.1111/ropr.12332

Schwab, K. (2016). *The fourth industrial revolution* (First U.S. ed.). Crown Business. https://law.unimelb.edu.au/__data/assets/pdf_file/0005/3385454/Schwab-The_Fourth_Industrial_Revolution_Klaus_S.pdf

Segun, S. T. (2021). Critically engaging the ethics of AI for a global audience. *Ethics and Information Technology*, *23*(2), 99–105. https://doi.org/10.1007/s10676-020-09570-y

Shakir, M., Marie-Therese, P., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, *33*(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Siafakas, N. M. (2021). Do we need a hippocratic oath for artificial intelligence scientists? *AI Magazine*, *42*(4), 57–61. https://doi.org/10.1609/aaai.12022

Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: Ethics of AI and ethical AI. *Journal of Database Management (JDM)*, *31*(2), 74–87. https://doi.org/10.4018/JDM.2020040105

Simanowski, R., Brodmer, M., & Chase, J. (2019). On the ethics of algorithmic intelligence. *Social Research*, *86*(2), 423–447. https://doi.org/10.1353/sor.2019.0021

Sison, A. J. G., & Redín, D. M. (2021). A neo-Aristotelian perspective on the need for artificial moral agents (AMAs). *Ai & Society*, *38*, 47–65. https://doi.org/10.1007/s00146-021-01283-0

Smith, N., & Vickers, D. (2021). Statistically responsible artificial intelligences. *Ethics and Information Technology*, *23*(3), 483–493. https://doi.org/10.1007/s10676-021-09591-1

Stanciu, V., & Rîndaşu, S.-M. (2021). Artificial intelligence in retail: Benefits and risks associated with mobile shopping applications. *Amfiteatru Economic*, *23*(56), 46–64. https://doi.org/10.24818/EA/2021/56/46

Stefan, R., & Carutasu, G. (2020). Explainable machine learning for ethical artificial intelligence based decisions. *Journal of Information Systems & Operations Management*, *14.1*, 151–161.

Taddeo, M., McNeish, D., Blanchard, A., & Edgar, E. (2021). Ethical principles for artificial intelligence in national defence. *Philosophy & Technology*, *34*(4), 1707–1729. https://doi.org/10.1007/s13347-021-00482-3

Tekin, Ş. (2021). Is big data the new stethoscope? Perils of digital phenotyping to address mental illness. *Philosophy & Technology*, *34*(3), 447–461. https://doi.org/10.1007/s13347-020-00395-7

Terblanche, N. (2020). A design framework to create artificial intelligence coaches. *International Journal of Evidence Based Coaching & Mentoring*, *18*(2), 152–165.

Thiebes, S., Lins, S., & Sunyaev, A. (2021). Trustworthy artificial intelligence. *Electronic Markets*, *31*(2), 447–464. https://doi.org/10.1007/s12525-020-00441-4

Thompson, C. L., & Morgan, H. M. (2020). Ethical barriers to artificial intelligence in the national health service, United Kingdom of Great Britain and Northern Ireland. *World Health Organization. Bulletin of the World Health Organization*, *98*(4), 293–295. https://doi.org/10.2471/BLT.19.237230

Tollon, F. (2021). Oxford handbook of ethics of AI. *Prometheus*, *37*(2), 211–215. International Bibliography of the Social Sciences (IBSS)

Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. *Journal of Business Research*, *122*, 502–517. https://doi.org/10.1016/j.jbusres.2020.09.009

Trunk, A., Birkel, H., & Hartmann, E. (2020). On the current state of combining human and artificial intelligence for strategic organizational decision making. *Business Research*, *13*(3), 875–919. https://doi.org/10.1007/s40685-020-00133-x

Vakkuri, V., Kemell, K.-K., Jantunen, M., Halme, E., & Abrahamsson, P. (2021). ECCOLA—A method for implementing ethically aligned AI systems. *Journal of Systems & Software*, *182*, 111067. https://doi.org/10.1016/j.jss.2021.111067

Votto, A. M., Valecha, R., Najafirad, P., & Rao, H. R. (2021). Artificial intelligence in tactical human resource management: A systematic literature review. *International Journal of Information Management Data Insights*, *1*(2), 100047. https://doi.org/10.1016/j.jjimei.2021.100047

Wallach, W. (2010). Robot minds and human ethics: The need for a comprehensive model of moral decision making. *Ethics and Information Technology*, *12*, 243–250.

Wamba, S. F., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, *164*, 120482. https://doi.org/10.1016/j.techfore.2020.120482

Wasilow, S., & Thorpe, J. B. (2019). Artificial intelligence, robotics, ethics, and the military: A Canadian perspective. *AI Magazine*, *40*(1), 37–48. https://doi.org/10.1609/aimag.v40i1.2848

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. *Business Horizons*, *64*(4), 465–474. https://doi.org/10.1016/j.bushor.2021.02.011

Weber, R. H. (2020). Socio-ethical values and legal rules on automated platforms: The quest for a symbiotic relationship. *Computer Law & Security Review*, *36*, 105380. https://doi.org/10.1016/j.clsr.2019.105380

White, C. L., & Boatwright, B. (2020). Social media ethics in the data economy: Issues of social responsibility for using Facebook for public relations. *Public Relations Review*, *46*(5), 101980. https://doi.org/10.1016/j.pubrev.2020.101980

Williams, N. L., Ferdinand, N., & Bustard, J. (2020). From WOM to aWOM—The evolution of unpaid influence: A perspective article. *Tourism Review of AIEST - International Association of Scientific Experts in Tourism*, *75*(1), 314–318. https://doi.org/10.1108/TR-05-2019-0171

Winkle, T., Erbsmehl, C., & Bengler, K. (2018). Area-wide real-world test scenarios of poor visibility for safe development of automated vehicles. *European Transport Research Review*, *10*(2), 1–15. https://doi.org/10.1186/s12544-018-0304-x

World Economic Forum. (2023). The future of jobs report 2023. https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf

Yam, J., & Skorburg, J. A. (2021). From human resources to human rights: Impact assessments for hiring algorithms. *Ethics and Information Technology*, *23*(4), 611–623. https://doi.org/10.1007/s10676-021-09599-7

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zuiderwijk, A., Chen, Y., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, *38*, 101577. https://doi.org/10.1016/j.giq.2021.101577

How to cite this article: Giarmoleo, F. V., Ferrero, I., Rocchi, M., & Pellegrini, M. M. (2024). What ethics can say on artificial intelligence: Insights from a systematic literature review. *Business and Society Review*, 1–35. https://doi.org/10.1111/basr.12336