IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

Artificial Intelligence of Ethics

Tushar Pawar[1], Rutwik Ghule[2], Shubham Karande[3], Prof. Y. G. Suryavanshi[4]
Student, Department of Computer Engineering[123]
Professor, Dept, of Computer Engineering[4]
Adsul Technical Campus, Chas, Ahilyanagar, Maharashtra, India

Abstract: *Artificial intelligence (AI) has profoundly changed and will continue to change our lives. AI is being applied in more and more fields and scenarios such as autonomous driving, medical care, media, finance, industrial robots, and internet services. The widespread application of AI and its deep integration with the economy and society have improved efficiency and produced benefits. At the same time, it will inevitably impact the existing social order and raise ethical concerns. Ethical issues, such as privacy leakage, discrimination, unemployment, and security risks, brought about by AI systems have caused great trouble to people. Therefore, AI ethics, which is a field related to the study of ethical issues in AI, has become not only an important research topic in academia, but also an important topic of common concern for individuals, organizations, countries, and society. This article will give a comprehensive overview of this field by summarizing and analyzing the ethical risks and issues raised by AI, ethical guidelines and principles issued by different organizations, approaches for addressing ethical issues in AI, and methods for evaluating the ethics of AI. Additionally, challenges in implementing ethics in AI and some future perspectives are pointed out. We hope our work will provide a systematic and comprehensive overview of AI ethics for researchers and practitioners in this field, especially the beginners of this research discipline. Impact Statement—AI ethics is an important emerging topic among academia, industry, government, society, and individuals. In the past decades, many efforts have been made to study the ethical issues in AI. This article offers a comprehensive overview of the AI ethics field, including a summary and analysis of AI ethical issues ethical guidelines and principles, approaches to address AI ethical issues, and methods to evaluate the ethics of AI technologies. Additionally, research challenges and future perspectives are discussed. This article will help researchers to gain a birds eye view of AI ethics, and thus facilitate their further investigation and research of AI.*

Keywords: Artificial intelligence (AI), AI ethics, ethical issue, ethical theory, ethical principle

I. INTRODUCTION

ARTIFICIAL intelligence (AI) [1] has achieved rapid and remarkable development during the last decade. AI technologies such as machine learning (ML), natural language processing, and computer vision are increasingly permeating and spreading to various disciplines and aspects of our society. AI is increasingly taking over human tasks and replacing human decision-making. It has been widely used in a variety of sectors, such as business, logistics, manufacturing, transportation, healthcare, education, state governance, etc. The application of AI has brought about efficiency improvement and cost reduction, which are beneficial for economic growth, social development, and human well-being [2]. For instance, the AI chatbot can respond to clients' inquiries at any time, which will improve the customers' satisfaction and the company's sales [3]. AI allows doctors to serve patients in remote locations through telemedicine services [4]. Itis no doubt that the rapid development and wide application of AI are already affecting our daily life, humanity, and society. However, at the same time, AI also poses many significant ethical risks or issues for users, developers, humans, and society. Over the past few years, many cases in which AI produced poor outcomes have been observed. For instance, in 2016, the driver of an electric Tesla car was killed in a road accident after its Autopilot mode failed to recognize an oncoming lorry [5]. Microsoft's AI chatting bot, Tay. ai, was taken down because it became racist and sexist only less than a day after she joined Twitter [6]. There are many other examples concerned with the failure, fairness, bias, privacy, and other ethical issues of AI systems [7]. More seriously, AI technology has begun to be used by criminals to harm others or the society. For example, criminals used AI-based software to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, December 2025 Impact Factor: 7.67

impersonate a chief executive's voice and demand a fraudulent transfer of $243 000 [8]. Therefore, it is urgent and critical to address the ethical issues or risks of AI so that AI can be built, applied, and developed ethically.

II. SCOPE AND METHODOLOGY

In this section, we first clarify the aspects and topics covered in this review and the links between these topics. Then, we describe the methodology followed in conducting this survey, including the literature search strategy and selection criteria.

A. Scope

The scope and topics of this article is described as follows. Investigation of ethical issues and risks of AI is the starting point of this review, since it is because of the existence of ethical issues in AI that the research field of AI ethics exists. Thus, it is necessary and important to clarify and understand the ethical problems existed in AI. Then, the ethical guidelines and principles, which direct the development and use of AI, are reviewed. As the ethical issues of AI have attracted more and more attention from various sectors of our society, many organizations (including academia, industry, and governments) have begun to discuss and seek possible frameworks, guidelines, and principles for solving AI ethics issues. These guidelines and principles provide valuable directions for practicing ethical AI. After clarifying the existing ethical issues and guidelines, we review the approaches to solving the ethical issues in AI.

Fig. 1. Topics covered in this article and the links between them.

B. Methodology

This review covers a wide variety of documents, including academic, organizational, government grey literature sources, and news report. The search of relevant literature was conducted in two phases. In the first phase, the entries or keywords that reflect different terms related to AI ethics are used to search on Google Scholar, Web of Science, IEEE Xplore, ACM Digital Library, Science Direct, Springer Link, are X iv, and Google. The entries or keywords used include: (ethics, ethical, responsibility, responsible, trustworthiness, trustworthy, transparent, explainable, fair, beneficial, robust, safe, private, sustainable) AND/OR (issues, risks, guideline, principle, approach, method, evaluation, assessment, challenge) AND (artificial intelligence, AI, machine learning, ML, intelligent system, intelligent agent). We mainly consider the literature published or released since 2010 and included as many related keywords as possible in titles. In the second phase, we checked the related work of literature found in the first phase, such as the cited articles and other work by the same authors of phase one. As for the ethical AI guidelines, we only collected these documents in English (or with official English translations) and can be visited or downloaded on the internet. A full list with URL links of collected ethical AI guidelines is provided in the Supplementary Materials of this article.

III. ETHICAL ISSUES AND RISKS OF AI

To address the ethical problems of AI, we must first recognize and understand the potential ethical issues or risks that AI may bring. Then, the necessary AI ethical guidelines, policies, principles, rules (i.e., Ethics of AI) can be formulated

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, December 2025 Impact Factor: 7.67

appropriately. With the adequate ethics of AI, we can design and build AI that behaves ethically (i.e., Ethical AI) [8]. The ethical issue of AI generally refers to the morally bad things or problematic outcomes relevant to AI (i.e., these issues and risks that are raised by the development, deployment, and use of AI) that need to be addressed. Many ethical issues, such as lack of transparency, privacy and accountability, bias and discrimination, safety and security problems, the potential for criminal and malicious use, and so on, have been identified from the applications and studies. This section focuses on ethical issues and risks of AI. First, four different categorizations of AI ethical issues in the literature are reviewed in Section III-A. Since these four categorizations either ignore some ethical issues or are too complicated to understand, we proposed a new categorization that classifies AI ethical issues into individual, societal, and environmental levels in Section III-B.

Our proposed categorization comprehensively covers the existing ethical issues and is easy to understand, which is helpful for understanding and analyzing the ethical problems caused by AI. Besides, we attempt to map the ethical issues associated with the stages of AI system's lifecycle in Section III-C. This would be beneficial for figuring out these issues during the AI system development process. The main goal of this section is to discuss and clarify the ethical issues of AI so that practitioners can recognize and understand these issues, and then help them to further study how to address AI ethical issues. The main contribution in this section is that we proposed a new categorization of AI ethical issues, which covers the ethical issues discussed in a clear and easy-to-understand manner. Additionally, the ethical issues associated with the stages of AI system's lifecycle is discussed.

IV. ETHICAL GUIDELINES AND PRINCIPLES FOR AI

As the ethical issues of AI have received more and more attention and discussions from various sectors of society, many organizations (including academia, industry, and government) have begun to discuss and seek the possible frameworks, guidelines and principles for solving AI ethics issues [78]. These guidelines and principles provide useful directions for practicing ethical AI. This section is dedicated to giving an up-to-date global landscape of the AI ethics guidelines and principles, which is achieved through the investigation of 146 reports, guidelines and recommendations related to AI ethics released by companies, organizations, and governments around the world since 2015. These guidelines and principles provide high-level guidance for the planning, development, production, and usage of AI and directions for addressing AI ethical issues.

A. Guidelines for AI Ethics

An excellent survey and analysis of the current principles and guidelines on ethical AI has been given in 2019 by Jobin et al. [12], who conducted a review of 84 ethical guidelines released by national or international organizations from various countries. Jobin et al. [12] found strong widespread agreement on five key principles, that is, transparency, justice and fairness, nonmaleficence, responsibility, and privacy, among many. However, many new guidelines and recommendations for AI ethics have been released in the past two years, making Jobin's paper obsolete because many important documents were not included. For instance, on November 24, 2021, UNESCO (the United Nations Educational, Scientific and Cultural Organization) adopted the Recommendation on the Ethics of Artificial Intelligence, which is the first ever global agreement on the ethics of AI [79]. To update and enrich the investigation on ethical AI guidelines and principles, based on the table of ethics guidelines for AI given in Jobin's paper [12] (only included 84 documents), we have collected many newly released AI ethical guidelines that are not included in Jobin's review. Finally, a total of 146 AI ethics guidelines have been collected.

A list of all the collected guidelines or documents is given in Table V of the Supplementary Materials. The number of guidelines issued each year from 2015 to 2021 is counted and listed in Table III. Itis apparent that the majority of the guidelines are released in the last five years, i.e., from 2016 to 2020. The number of guides published in 2018 was the largest, with 53, accounting for 36.3% of the total number. Additionally, the number of AI guidelines issued by each country is listed in Table IV. Furthermore, the percentages of guidelines released by different types of issuers (including government, industry, academia, and other organizations) are shown in Fig. 3. It can be seen from Fig. 3 that governments, companies, and academia all have shown strong concerns about AI ethics.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

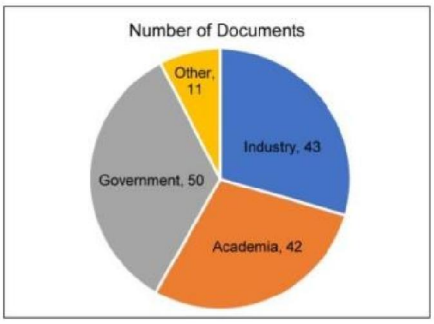

Fig. 3. Percentage of guidelines released by different types of issuers.

B. Principles for AI Ethics

The ethical principles that are featured in the collected 146 guidelines are listed in Table I of the Supplementary Materials. According to the table, there is an obvious convergence emerging round five important ethical principles: transparency, fairness and justice, responsibility, nonmaleficence, and privacy. The 11 ethical principles identified in the existing AI guidelines are described and explained in the following.

Transparency: Transparency is one of the most widely discussed principles in the AI ethics debate. The transparency of AI mainly involves the transparency of the AI technology itself, and the transparency of the developing and adopting of the AI [13]. On one hand, transparency of AI involves the interpretability of a given AI system, that is, the ability to know how and why a model performed the way it did in a specific context and thus to understand the rationale behind its decision or behavior. This aspect of transparency is usually mentioned as the metaphor of "opening the black box of AI." It concerns interpretability, explainability, or understandability. On the other hand, transparency of AI includes the justifiability or rationality of the design and implementation process of the AI system and that of its outcome. In other words, the design and implementation process of the AI system and its decision or behavior must be justifiable and visible.

Fairness & Justice: The principle of justice and fairness states that the development, deployment, and use of AI must be just and fair so that the AI system should not result in discriminations or bias against individuals, communities, or groups [80]. Discrimination and unfair outcomes brought by AI algorithms have become a hot topic in the media and academia. Consequently, fairness and justice principle has attracted considerable attention during the last few years

.3) Responsibility and Accountability: The principle of responsibility and accountability requires that AI must be auditable, that is, the designers, developers, owners, and operators of AI are responsible and accountable for an AI system's behaviors or decisions, and are therefore considered responsible for harms or bad outcomes it might cause [51]. The designers, builders, and users of AI systems are stakeholders in the moral or ethical implications of their use, misuse, and behavior, and they have the responsibility and opportunity to shape these implications. This requires that appropriate mechanisms should be established to ensure responsibility and accountability for AI systems and their results, both before and after their development, deployment, and use 4) Nonmaleficence: The nonmaleficence basically means to do no harm or avoid imposing risks of harm to others [81], [82]. Thus, the nonmaleficence principle of AI generally refers to that AI systems should not cause or exacerbate harm to humans or adversely affect human beings. This entails the protection of human dignity as well as mental and physical integrity.The nonmaleficence principle requires that AI systems and the environments in which they operate must be safe and secure so that they are not open to malicious use. With some of the fatal accidents coming from autonomous cars and robots, avoiding harm to human beings is one of the greatest concerns in AI ethics. Hence, most of the ethical guidelines put a strong emphasis on ensuring no harm to human beings through the safety and security of AI.

Privacy: The privacy principle aims to ensure respect for privacy and data protection when using AI systems. AI systems should preserve and respect privacy rights and data protection as well as maintain data security. This involves providing effective data governance and management for all data used and generated by the AI system throughout its entire lifecycle [83]. Specifically, data collection, usage and storage must comply with laws and regulations related to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, December 2025 Impact Factor: 7.67

privacy and data protection. Data and algorithms must be protected against theft. Once information leakage occurs, employers or AI providers need to inform employees, customers, partners, and other relevant individuals as soon as possible to minimize the loss or impact caused by the leakage.

Beneficence: The principle of beneficence states that AI shall do people good and benefit humanity [82]. This principle indicates that AI technology should be used to bring beneficial outcome and impactto individuals, society, and the environment [84].

When developing an AI system, its objectives should be clearly defined and justified. The use of AI technology to help address global concerns should be encouraged, such as using AI to help us to handle food security, pollution, and contagion like AIDS and COVID 19.

Freedom and Autonomy: Freedom and autonomy, which generally refers to the ability of a person to make decisions respect to his goals and wishes, is the core value for citizens in democratic societies. Therefore, it is important that the use of AI does not harm or encumber the freedom and autonomy for us. When we apply AI agents, we are willing to give up part of our decision-making authority to AI machines. Thus, upholding the principle of freedom and autonomy in the context of AI means to strike a balance between the decision-making power we maintain for ourselves and that which we cede toAI [84].

Solidarity: The solidarity principle entails that the development and application of an AI system must be compatible with maintaining the bounds of solidarity among people and generations. In other words, AI should promote social security and cohesion, and should not jeopardize social bonds and relationships [13].

Sustainability: Due to climate change and ongoing environmental damage, the importance of sustainability has received more and more attention. Like other fields and disciplines, AI is affected and needs to be included in the sustainable development agenda. The sustainability principle represents that the production, management, and implementation of AI must be sustainable and avoid environmental harm. In other words, AI technology must meet the requirements of ensuring the continued prosperity of mankind and preserving a good environment for future generations [85]. AI systems promise to help tackling some of the most pressing societal concerns, but it must be ensured that this happens in the most environmentally friendly way possible.

Trust: Trustworthiness is a prerequisite for people and societies to adopt AI, since trust is a basic principle for interpersonal interactions and social operation. The trust in the development, deployment and use of AI systems is not only related to the inherent characteristics of the technology, but also related to the quality of the socio-technical system involving AI applications. Therefore, moving toward trustworthy AI not only concerns the trustworthiness of the AI system itself, but also requires a holistic and systematic approach that covers the trustworthiness of all participants and processes that are the entire life cycle of the system [86].

Dignity: Human dignity encompasses the belief that all people possess an intrinsic value that is tied solely to their humanity, i.e. ,it has nothing to do with their class, race, gender, religion, abilities, or any other factor other than them being human, and this intrinsic value should never be diminished, compromised, or repressed by other people nor by technologies like AI. It is important that AI should not infringe or harm the dignity of end-users or other members of society. As a result, respecting human dignity is an important principle that should be considered in AI ethics. AI system should hence be developed in a way that respects, supports, and protects people's physical and mental integrity, personal and cultural sense of identity, and satisfaction of their basic needs [13].

V. APPROACHES TO ADDRESS ETHICAL ISSUES IN AI

This section reviews the approaches to address or mitigate ethical issues of AI. As AI ethics is a broad and multidisciplinary field, we attempt to provide a comprehensive overview of the existing and potential approaches for addressing AI ethical issues, including ethical, technological, and legal approaches, rather than solely focusing on technological approaches that are of interest to the field of AI/ML community. This review of multidisciplinary approaches for addressing AI ethical problems not only provides an informative summary about the approaches to ethical AI but also suggests the researchers in AI community to seek solutions to AI ethical issues from a variety of perspectives rather than relying solely on technological approaches. As AI ethical issues are complex with

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, December 2025 Impact Factor: 7.67

multidisciplinary problems, it may be possible to solve these problems effectively only through the cooperation of different methods.

A Challenges in Evaluating Ethics in AI

Ethics is inherently a qualitative concept that depends on many features that are hard to quantify, e.g., culturally or racially related features. Hence, it is very hard, if not impossible, to define ethics precisely.

As a result, the evaluation of AI ethics will always have some subjective elements, depending on the people who are assessing AI. This poses challenges to the research and applications of AI ethics.

B. Future Perspectives

In this section, some future perspectives are pointed out, which may be valuable for future research. First, for implementing ethics in AI, it should be pointed out that humans never use only one single ethical theory, but will switch between different theories according to the situation or context they are facing [134]. This is not only because human beings are not purely rational agents that economic theory wants us to believe, but also because strict adherence to any moral theory can lead to undesirable results. This means that AI systems should be provided with representations of different ethical theories and the ability to choose between these ethical theories. Here we call this multi-theory approach. In multi theory approach, AI systems can interchangeably apply different theories depending on the type of situation. Furthermore, the combination of normative ethical theories and domain-specific ethics which accepted by domain experts is worthy of implementing since an ethical AI system need to be accepted by its users.

VI. CONCLUSION

Based on our review of AI ethics and the many complexities and challenges described in this article, it is clear that attempting to address ethical issues in AI and to design ethical AI systems that are able to behave ethically is a tricky and complex task. However, whether AI can play an increasingly important role in our future society largely depends on the success of ethical AI systems. The discipline of AI ethics requires a joint effort of AI scientists, engineers, philosophers, users, and government policymakers. This article provides a comprehensive overview of AI ethics by summarizing and analyzing the ethical risks and issues raised by AI, ethical guidelines and principles issued by different organizations, approaches for addressing ethical issues in AI or fulfilling ethical principles of AI, and methods for evaluating the ethics (or morality) of AI. Furthermore, some challenges in the practice of AI ethics and some future research directions are pointed out. However, AI ethics is a very broad and multidisciplinary research area. It is impossible to cover all possible topics in this area with one review article. We hope this article can serve as a starting point for people who are interested in AI ethics to gain a sufficient background and a bird's eye view so that further investigation can be pursued by them.

REFERENCES

[1]. D. Roselli, J. Matthews, and N. Talagala, "Managing bias in AI," in Proc. World Wide Web Conf., 2019, pp. 539–544.

[2]. Y. Gorodnichenko, T. Pham, and O. Talavera, "Social media, sentiment and public opinions: Evidence from #Brexit and #USElection," Eur. Econ.Rev., vol. 136, Jul. 2021, Art. no. 103772.

[3]. N. Thurman, "Making 'The daily me': Technology, economics and habit in the mainstream assimilation of personalized news," Journalism,vol. 12, no. 4, pp. 395–415, 2011.

[4]. J. Donath, "Ethical issues in our relationship with artificial entities," in The Oxford Handbook of Ethics of AI. M. D. Dubber, F. Pasquale, and S.Das, Eds., Oxford, U.K.: Oxford Univ. Press, 2020, pp. 51–73.

[5]. E. Magrani, "New perspectives on ethics and the laws of artificial intelligence," Internet Policy Rev., vol. 8, 2019, Art. no. 3.

[6]. M. P. Wellman and U. Rajan, "Ethical issues for autonomous tradingagents," Minds Mach., vol. 27, no. 4, pp. 609–624, 2017.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, December 2025 Impact Factor: 7.67

[7]. U. Pagallo, "The impact of AI on criminal law, and its two fold procedures," in Research Handbook on the Law of Artificial Intelligence,W. Barfield and U. Pagallo, Eds., Cheltenham U.K.: Edward Elgar Publishing, 2018, pp. 385–409.

[8]. E. Dacoronia, "Tort law and new technologies," in Legal Challenges in the New Digital Age, A. M. López Rodríguez, M. D. Green, and M.L. Kubica, Eds., Leiden, The Netherlands: Koninklijke Brill NV, 2021, pp. 3–12.

[9]. J. Khakurel, B. Penzenstadler, J. Porras, A. Knutas, and W. Zhang, "The rise of artificial intelligence under the lens of sustainability," Technologies, vol. 6, no. 4, 2018, Art. no. 100.

[10]. S. Herat, "Sustainable management of electronic waste (e-Waste)," Clean Soil Air Water, vol. 35, no. 4, pp. 305–310, 2007.

[11]. E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in NLP," in Proc. 57th Annu. Meeting Assoc. Comput, Linguistics, 2019, pp. 3645–3650.

[12]. V. Dignum, "Ethics in artificial intelligence: Introduction to the special issue," Ethics Inf. Technol., vol. 20, no. 1, pp. 1–3, 2018.

[13]. S. Corbett-Davies, E. Pierson, A. Feller, S. Goel, and A. Huq, "Algorithmic decision making and the cost of fairness," in Proc. 23rd ACMSIGKDD Int. Conf. Knowl. Discov. Data Mining, 2017, pp. 797–806.

[14]. R. Caplan, J. Donovan, L. Hanson, and J. Matthews, "Algorithmic accountability: A primer," Data Soc., vol. 18, pp. 1–13, 2018.

[15]. R. V. Yampolskiy, "On controllability of AI," Jul. 2020. [Online]. Available: https://arxiv.org/pdf/2008.04071

[16]. B. C. Stahl, J. Timmermans, and C. Flick, "Ethics of emerging information and communication technologies," Sci. Public Policy, vol. 44, no. 3, pp. 369–381, 2017.

[17]. L. Vesnic-Alujevic, S. Nascimento, and A. Pólvora, "Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks," Telecommun. Policy, vol. 44, no. 6, 2020,Art. no. 101961.

[18]. U. G. Assembly, "Universal declaration of human rights," UN Gen Assem., vol. 302, no. 2, pp. 14–25, 1948.

[19]. K. Arkoudas, S. Bringsjord, and P. Bello, "Toward ethical robots via mechanized deontic logic," in Proc. AAAI Fall Symp. Mach. Ethics, 2005,pp. 17–23.

[20]. S. Bringsjord and J. Taylor, "Introducing divine-command robot ethics," in Robot Ethics: The Ethical and Social Implication of Robotics. 2012,pp. 85–108.

[21]. N. S. Govindarajulu and S. Bringsjord, "On automating the doctrine of double effect," in Proc. 26th Int. Joint Conf. Artif. Intell., 2017, pp. 4722–4730.

[22]. F. Berreby, G. Bourgne, and J.-G. Ganascia, "A declarative modular framework for representing and applying ethical principles," in Proc. 16th Conf. Auton. Agents MultiAgent Syst., 2017, pp. 96–104.

[23]. V. Bonnemains, C. Saurel, and C. Tessier, "Embedded ethics: Some echnical and ethical challenges," Ethics Inf. Technol., vol. 20, no. 1, pp. 41–58, 2018.

[24]. G. S. Reed, M. D. Petty, N. J. Jones, A. W. Morris, J. P. Ballenger, and H. S. Delugach, "A principles-based model of ethical considerations in military decision making," J. Defense Model. Simul., vol. 13, no. 2,pp. 195–211, 2016.

[25]. L. Dennis, M. Fisher, M.Slavkovik, and M.Webster, "Formal verificationmof ethical choices in autonomous systems," Robot. Auton. Syst., vol. 77, pp. 1–14, 2016.

[26]. A R. Honarvar and N. Ghasem-Aghaee, "Casuist BDI-Agent: A new extended BDI architecture with the capability of ethical reasoning," in Proc. Int. Conf. Artif. Intell. Comput. Intell., 2009, pp. 86–95.

[27]. S. Rao and M. P. Georgeff, "BDI agents: From theory to practice," in Proc. 1st Int. Conf. Multiagent Syst., 1995, pp. 312–319.

[28]. S. Armstrong, "Motivated value selection for artificial agents," in Proc. AAAI Workshop Artif. Intell. Ethics, Jan. 2015, pp. 12–20.

[29]. U. Furbach, C. Schon, and F. Stolzenburg, "Automated reasoning in deontic logic," in Proc. 8th Int. Workshop Multi-Disciplinary Trends Artif. Intell., 2014, pp. 57–68.

DOI: 10.48175/IJARSCT-30166

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, December 2025
Impact Factor: 7.67

[30]. D. Howard and I. Muntean, "Artificial moral cognition: Moral functionalism and autonomous moral agency," in Philosophical Studies Series, Philosophy and Computing, T. M. Powers, Ed. Cham, Switzerland: Springer, 2017, pp. 121–159.