GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

# An Ethical Review of Human Existence in the Era of Artificial Intelligence: Automation, Autonomous Weapons, and the Threat of Superintelligence

**Wahyu Kusuma Putra[1] , Evy Nurmiati[2],**
[1,2]Universitas Islam Negri Syarif Hidayatullah Jakarta, Indonesia
[1]wahyu.kusuma24@mhs.uinjkt.ac.id, [2]evy.nurmiati@uinjkt.ac.id,

open access

***Corresponding Author**

**Article History:**
Submitted: 1 July 2026
Accepted: 6 July 2026
**Published: 7 July 2026**

**Keywords:**
Artificial Intelligence; Automation; Human Existence; Human-Centered Ethical Governance; Superintelligence.



**ABSTRACT**

Artificial Intelligence (AI) has rapidly evolved into one of the most influential technologies of the twenty-first century, transforming human activities across economic, social, political, and security domains while simultaneously raising increasingly complex ethical concerns regarding human existence. This study aims to critically examine the ethical implications of AI by focusing on three interconnected domains automation, autonomous weapons, and superintelligence and to identify ethical principles capable of supporting responsible AI governance. A Critical Literature Review (CLR) was employed to synthesize peer-reviewed journal articles, conference proceedings, and internationally recognized policy documents retrieved from major academic databases and global organizations, including Scopus, IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library, UNESCO, OECD, the European Commission, and the International Committee of the Red Cross. The findings reveal that although AI substantially enhances productivity, innovation, and decision-making capabilities, it also introduces significant ethical challenges, including workforce transformation, accountability gaps in autonomous military operations, diminished human autonomy in AI-assisted decision-making, and long-term existential risks associated with highly autonomous intelligent systems. The synthesis further demonstrates that current ethical approaches remain fragmented because they frequently address these issues independently rather than as interconnected dimensions of AI development. This study concludes that a Human-Centered Ethical Governance Framework, grounded in the principles of human dignity, human autonomy, transparency, accountability, fairness, and meaningful human oversight, provides a more comprehensive approach for governing AI technologies and ensuring that future AI development remains aligned with fundamental human values while supporting sustainable coexistence between humanity and intelligent machines.

## INTRODUCTION

Artificial Intelligence (AI) has become one of the defining technologies of the twenty-first century, fundamentally transforming the relationship between humans and intelligent machines. Unlike previous technological revolutions that primarily enhanced physical productivity, contemporary AI increasingly performs cognitive functions such as learning, reasoning, prediction, language generation, and autonomous decision-making. Consequently, AI is no longer merely regarded as a technological instrument but as a transformative force capable of reshaping economic systems, social interactions, governance structures, and even the philosophical understanding of what it means to be human (Russell, 2019; UNESCO, 2021; OECD, 2024).

The growing integration of AI into everyday life has generated remarkable benefits across numerous sectors. Intelligent systems have improved medical diagnosis, optimized industrial production, enhanced scientific discovery, supported educational innovation, strengthened financial analysis, and accelerated digital transformation worldwide. Recent advances in foundation models and generative AI further demonstrate that machines are becoming increasingly capable of performing tasks previously considered exclusive to human intelligence (European Commission, 2019; Stanford HAI, 2025). These developments position AI as one of the primary drivers of technological progress and economic competitiveness in the digital era.

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

Beyond improving productivity, intelligent digital platforms have also transformed the way users interact with technology and digital content. Recent studies indicate that user engagement is influenced not only by technological capabilities but also by platform features that facilitate interaction, participation, and meaningful communication. These findings demonstrate that technological innovation should be evaluated from both technical and human interaction perspectives, particularly as AI becomes increasingly integrated into digital ecosystems (Subiyakto, A., Yudhanta, S., Nurmiati, E., et al. 2025).

Despite these significant achievements, the rapid advancement of AI has also transformed ethical discussions surrounding technology. Earlier debates primarily focused on algorithmic efficiency, computational accuracy, and technological performance. Today, however, ethical concerns extend far beyond technical capabilities toward broader questions regarding human dignity, autonomy, justice, accountability, transparency, and the long-term preservation of human agency. International organizations increasingly recognize that technological innovation should not be evaluated solely according to economic efficiency but also according to its ability to protect fundamental human rights and promote social well-being (UNESCO, 2021; OECD, 2024; European Commission, 2019).

Among the various ethical challenges associated with AI, three emerging domains have become particularly significant because of their direct implications for human existence. The first concerns automation, where intelligent systems increasingly replace both manual and cognitive labor. While automation contributes substantially to productivity and organizational efficiency, it simultaneously raises concerns regarding technological unemployment, workforce transformation, digital inequality, and the gradual reduction of meaningful human participation in decision-making processes (Acemoglu & Restrepo, 2020; International Labour Organization, 2023; Brynjolfsson & McAfee, 2014).

The second domain concerns Autonomous Weapons Systems (AWS), which represent one of the most controversial applications of AI. Unlike conventional military technologies, autonomous weapons possess varying levels of decision-making capability, enabling them to identify, select, and engage targets with limited human intervention. Such developments fundamentally challenge established principles of humanitarian law, legal accountability, proportionality, and moral responsibility. Consequently, international organizations and AI researchers increasingly advocate maintaining meaningful human control over lethal autonomous systems to ensure that ethical judgment remains inseparable from military decision-making (Sharkey, 2019; International Committee of the Red Cross, 2021; Russell, 2019).

The third and perhaps most profound challenge concerns the potential emergence of **superintelligence**. Although current AI systems have not yet reached superintelligent capabilities, rapid progress in machine learning has intensified discussions concerning AI alignment, loss of human control, and long-term existential risk. Bostrom (2014) argues that future AI systems exceeding human cognitive abilities could fundamentally alter civilization if their objectives become misaligned with human values. More recent studies similarly emphasize that governance, safety research, and international cooperation should be developed proactively rather than reactively to minimize future risks associated with highly autonomous AI (Russell, 2019; Hendrycks et al., 2023).

The ethical implications of these three domains demonstrate that AI should no longer be understood merely as a technological innovation but as a phenomenon capable of redefining the relationship between humans and intelligent systems. As AI assumes increasingly complex cognitive responsibilities, ethical governance becomes essential to preserve human dignity, democratic values, institutional accountability, and meaningful human autonomy. Consequently, AI ethics has evolved into an interdisciplinary field integrating perspectives from computer science, philosophy, law, political science, sociology, and public policy (Floridi & Cowls, 2019; Hagendorff, 2020).

Previous studies have made substantial contributions to understanding specific dimensions of AI ethics. Floridi and Cowls (2019) proposed a unified framework emphasizing beneficence, non-maleficence, autonomy, justice, and explicability as core principles of ethical AI. Hagendorff (2020) critically evaluated the gap between AI ethical principles and practical implementation, while Veale and Zuiderveen Borgesius (2021) analyzed regulatory challenges within the European Union's AI governance framework. Other researchers have investigated automation and labor-market transformation (Acemoglu & Restrepo, 2020), autonomous weapons and humanitarian law (Sharkey, 2019), as well as AI alignment and existential risk associated with advanced AI systems (Russell, 2019; Hendrycks et al., 2023). Collectively, these studies provide valuable insights into individual aspects of AI ethics.

Nevertheless, the existing literature remains largely fragmented. Most previous studies examine automation, autonomous weapons, or superintelligence independently, with relatively limited attention given to their interconnected ethical implications for human existence. Consequently, there remains a lack of comprehensive analysis integrating these three domains into a unified ethical perspective capable of explaining how increasing AI autonomy

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

collectively influences human dignity, autonomy, governance, and long-term societal sustainability (Stix, 2021; UNESCO, 2021).

Therefore, this study aims to critically examine the ethical implications of automation, autonomous weapons, and superintelligence through a Critical Literature Review (CLR). By synthesizing contemporary scholarly literature together with internationally recognized ethical frameworks including the UNESCO Recommendation on the Ethics of Artificial Intelligence, the OECD AI Principles, and the European Commission's Ethics Guidelines for Trustworthy AI this study seeks to identify recurring ethical principles and evaluate their implications for human existence.

Unlike previous studies that primarily discuss these issues separately, this article contributes by integrating automation, autonomous weapons, and superintelligence into a single analytical framework centered on human existence. Furthermore, this study proposes a Human-Centered Ethical Governance Framework that emphasizes human dignity, accountability, transparency, fairness, human autonomy, and meaningful human oversight as fundamental principles for ensuring that future AI development remains aligned with human values and contributes to the sustainable coexistence of humanity and increasingly intelligent technologies.

## LITERATURE REVIEW

### A. Artificial Intelligence and Human Existence

Artificial Intelligence (AI) has evolved from a computational tool designed to automate routine tasks into an autonomous technology capable of supporting complex cognitive processes. Contemporary AI systems are increasingly able to recognize patterns, learn from experience, generate knowledge, and perform sophisticated decision-making across multiple domains. This transformation has significantly influenced economic development, healthcare, education, manufacturing, transportation, and public administration, positioning AI as one of the primary drivers of the Fourth Industrial Revolution (Russell, 2019; UNESCO, 2021).

Recent studies argue that AI should no longer be viewed merely as a technological innovation but as a socio-technical system that reshapes human interaction, institutional governance, and societal values. Floridi and Cowls (2019) emphasize that AI increasingly affects moral decision-making and therefore requires ethical principles capable of protecting human dignity, autonomy, and justice. Similarly, UNESCO (2021) argues that AI development should promote human rights, cultural diversity, environmental sustainability, and social inclusion rather than focusing exclusively on technological efficiency.

These perspectives indicate a fundamental shift in AI research from technical optimization toward understanding AI's influence on human existence. Consequently, ethical governance has become an essential component of AI development to ensure that technological advancement remains aligned with fundamental human values.

### B. Ethical Foundations of Artificial Intelligence

Ethical discussions concerning AI have expanded significantly over the past decade. Early studies primarily addressed algorithmic performance and computational efficiency, whereas contemporary research increasingly focuses on the moral and societal consequences of autonomous intelligent systems. Floridi and Cowls (2019) proposed five ethical principles beneficence, non-maleficence, autonomy, justice, and explicability as a unified framework for evaluating AI technologies. These principles have become one of the most widely adopted conceptual foundations in AI ethics.

International organizations have further strengthened these ethical principles through global governance initiatives. UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021) identifies transparency, accountability, fairness, privacy, human dignity, and meaningful human oversight as essential requirements for responsible AI. Likewise, the OECD AI Principles (2024) emphasize inclusive growth, democratic values, robustness, transparency, and accountability as fundamental characteristics of trustworthy AI.

However, Hagendorff (2020) argues that the existence of ethical principles alone is insufficient because many AI systems continue to exhibit implementation gaps between ethical guidelines and real-world practice. This observation demonstrates that AI ethics should be understood not merely as a normative framework but also as a governance challenge requiring practical implementation across technological, legal, and institutional contexts.

### C. Ethical Challenges of Automation

Automation represents the most mature and widely adopted application of AI. Numerous studies acknowledge

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

that AI-driven automation significantly improves productivity, operational efficiency, and organizational competitiveness by replacing repetitive and analytical tasks (Brynjolfsson & McAfee, 2014; Acemoglu & Restrepo, 2020).

Despite these benefits, researchers consistently identify workforce transformation as automation's primary ethical challenge. Acemoglu and Restrepo (2020) demonstrate that industrial robot adoption significantly affects employment opportunities and wage distribution, particularly among routine occupations. Similarly, the International Labour Organization (2023) concludes that generative AI will reshape employment structures by automating cognitive tasks while simultaneously increasing demand for digital competencies and lifelong learning.

Beyond economic concerns, automation also raises ethical questions regarding human autonomy and dignity. Russell (2019) argues that increasing dependence on AI-supported decision-making may reduce meaningful human participation in organizational processes. Consequently, contemporary research increasingly recommends designing AI systems that augment rather than replace human capabilities, thereby preserving human agency within technologically advanced workplaces (UNESCO, 2021).

### D. Ethical Challenges of Autonomous Weapons

Among contemporary AI applications, Autonomous Weapons Systems (AWS) represent one of the most controversial ethical issues. These systems possess varying levels of autonomy that enable them to identify, select, and engage military targets with limited human intervention. While proponents argue that autonomous weapons may improve operational precision and reduce military casualties, critics emphasize their potential to undermine humanitarian law and ethical accountability (Sharkey, 2019).

The International Committee of the Red Cross (2021) highlights that autonomous weapons fundamentally challenge existing legal frameworks because responsibility becomes increasingly difficult to assign when lethal decisions involve autonomous systems. Russell (2019) similarly argues that delegating life-and-death decisions to AI undermines the moral accountability traditionally associated with human judgment.

Recent studies therefore advocate maintaining meaningful human control throughout the deployment of autonomous military technologies. Rather than relying solely on technological performance, ethical governance should prioritize legal responsibility, transparency, and international cooperation to prevent the uncontrolled proliferation of AI-enabled weapons (Future of Life Institute, 2017; UNESCO, 2021).

### E. Ethical Challenges of Superintelligence

Superintelligence represents one of the most debated long-term topics within AI ethics. Bostrom (2014) defines superintelligence as an AI system capable of substantially outperforming human intelligence across virtually all cognitive domains. Although such systems remain hypothetical, many researchers argue that governance and ethical preparation should begin before technological realization becomes imminent.

The central concern within superintelligence research is AI alignment the challenge of ensuring that increasingly autonomous AI systems continue pursuing objectives compatible with human values. Russell (2019) argues that future AI safety depends fundamentally upon maintaining uncertainty regarding human preferences and continuously aligning AI objectives with societal values.

Recent studies further identify existential risk as a defining characteristic distinguishing superintelligence from contemporary AI applications. Hendrycks et al. (2023) argue that advanced AI systems may fundamentally reshape geopolitical stability, institutional governance, and humanity's long-term future if appropriate governance mechanisms are not established. Consequently, researchers increasingly recommend proactive international cooperation, AI safety research, and responsible governance to minimize future existential risks.

### F. Research Gap

The reviewed literature demonstrates considerable progress in understanding individual dimensions of AI ethics. Existing studies comprehensively examine trustworthy AI, automation, autonomous weapons, AI alignment, and existential risk. However, these topics are predominantly investigated independently within separate disciplinary contexts.

This fragmentation limits understanding of how automation, autonomous weapons, and superintelligence collectively influence human existence. While previous studies contribute important theoretical and practical insights, relatively few provide an integrated ethical analysis connecting these three domains within a single conceptual framework centered on human dignity, autonomy, accountability, and long-term societal sustainability.

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

Accordingly, this study addresses this gap by synthesizing contemporary literature through a **Critical Literature Review (CLR)**. Rather than examining each ethical issue separately, the study integrates automation, autonomous weapons, and superintelligence into a unified perspective of **Human-Centered Ethical Governance**, providing a broader understanding of how increasingly autonomous AI systems reshape the ethical relationship between humanity and intelligent technologies.

## METHOD

### A. Research Design

This study employed a Critical Literature Review (CLR) to examine the ethical implications of Artificial Intelligence (AI) on human existence, with a particular focus on automation, autonomous weapons, and the potential emergence of superintelligence. Unlike a traditional narrative literature review, the CLR approach enables researchers to critically evaluate, compare, and synthesize findings from previous studies while identifying conceptual relationships, research gaps, and future research directions. This method is particularly suitable for ethical AI research because it integrates perspectives from computer science, philosophy, law, public policy, and social sciences to develop a comprehensive understanding of contemporary AI ethics (Grant & Booth, 2009; Snyder, 2019).

### B. Literature Search Strategy

The literature search was conducted systematically using internationally recognized academic databases, including Scopus, IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library, and Google Scholar. In addition, official policy documents and ethical guidelines published by UNESCO, the Organisation for Economic Co-operation and Development (OECD), the European Commission, the International Committee of the Red Cross (ICRC), the National Institute of Standards and Technology (NIST), and the World Economic Forum (WEF) were included to strengthen the governance perspective.

The search process employed combinations of the following keywords:

"Artificial Intelligence Ethics"
"Human Existence"
"Automation"
"Autonomous Weapons"
"Superintelligence"
"Human-Centered AI"
"AI Governance"
"Responsible Artificial Intelligence"
"AI Alignment"
"Existential Risk"

The search focused primarily on publications released between 2020 and 2025 to ensure that the analysis reflected the latest developments in AI ethics. Nevertheless, several seminal works published before 2020, such as those by Bostrom (2014), Brynjolfsson and McAfee (2014), and Russell (2019), were intentionally included because they constitute foundational references within AI ethics and governance.

### C. Inclusion and Exclusion Criteria

To ensure the quality and relevance of the reviewed literature, explicit inclusion and exclusion criteria were established.

Table 1. Inclusion and Exclusion Criteria

| Inclusion Criteria | Exclusion Criteria |
|---|---|
| Peer-reviewed journal articles | Non-academic blogs and opinion articles |
| Conference proceedings indexed by IEEE or ACM | News articles without scientific review |
| International reports from UNESCO, OECD, NIST, ICRC, and WEF | Publications lacking identifiable authorship |

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

| | |
|---|---|
| Publications discussing AI ethics, automation, autonomous weapons, superintelligence, AI governance, or human-centered AI | Publications unrelated to AI ethics |
| English-language publications | Non-English publications without verified translations |
| Publications from 2020–2025 (with selected seminal works before 2020) | Duplicate publications |

Following the application of these criteria, the selected literature consisted of high-quality academic publications and internationally recognized policy documents directly relevant to the objectives of this study.

### D. Data Analysis

The selected literature was analyzed using **Thematic Analysis**, allowing ethical issues to be systematically categorized according to recurring concepts and patterns identified across previous studies (Braun & Clarke, 2006). Four analytical stages were conducted.

The first stage involved familiarization with the literature through comprehensive reading of the selected publications. The second stage consisted of identifying recurring ethical themes and assigning initial codes related to human dignity, accountability, transparency, fairness, human autonomy, human oversight, AI alignment, and existential risk. The third stage compared findings across different studies to identify similarities, differences, and unresolved ethical debates. Finally, the synthesized findings were interpreted using internationally recognized ethical frameworks, including UNESCO's Recommendation on the Ethics of Artificial Intelligence, the OECD AI Principles, and the Ethics Guidelines for Trustworthy AI issued by the European Commission.

### E. Research Workflow

The overall research process followed five sequential stages, beginning with the identification of the research problem and ending with the synthesis of ethical findings into a conceptual framework. The workflow was designed to ensure that the literature selection, critical analysis, and synthesis were conducted systematically and transparently. Figure 1 illustrates the methodological workflow adopted in this study.

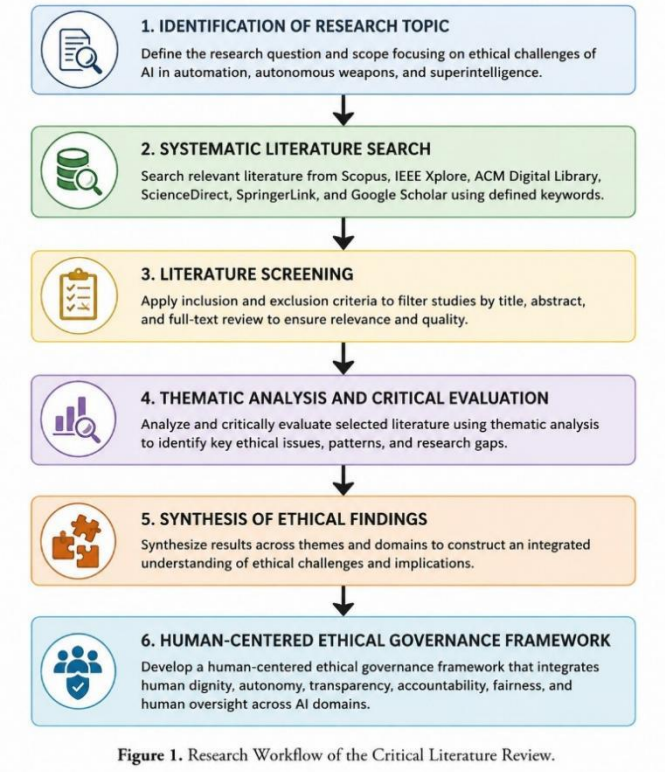


**Figure 1.** Research Workflow of the Critical Literature Review.

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

The workflow demonstrates that the study began by identifying the research problem, followed by a systematic search of relevant literature from international academic databases and institutional reports. The selected studies were then screened using predefined inclusion and exclusion criteria before being critically analyzed through thematic analysis. Finally, the findings were synthesized to develop the proposed **Human-Centered Ethical Governance Framework**, which serves as the conceptual contribution of this study.

## RESULT

The Critical Literature Review identified three major ethical domains that consistently emerge in contemporary discussions concerning the impact of Artificial Intelligence (AI) on human existence. Through thematic analysis, the reviewed literature demonstrates that ethical concerns evolve alongside the increasing autonomy of AI systems. Rather than focusing solely on technological performance, recent studies emphasize broader issues related to human dignity, accountability, autonomy, fairness, transparency, and meaningful human oversight.

The synthesis further indicates that these ethical concerns can be grouped into three interconnected domains: **automation**, **autonomous weapons**, and **superintelligence**. Although each domain presents distinct ethical implications, they collectively illustrate the progressive transformation of AI from a productivity-enhancing technology into a governance challenge that influences social, legal, political, and existential dimensions of human life.

Table 2. Major Ethical Domains Identified Through the Critical Literature

| Ethical Domain | Primary Ethical Concern | Main Impact on Human Existence | Representative References |
|---|---|---|---|
| Automation | Workforce transformation | Employment, inequality, human autonomy | Acemoglu & Restrepo (2020); ILO (2023) |
| Autonomous Weapons | Accountability | Humanitarian law and moral responsibility | Sharkey (2019); ICRC (2021) |
| Superintelligence | AI alignment | Existential risk and long-term governance | Russell (2019); Hendrycks et al. (2023) |

Table 2 summarizes the three principal ethical domains identified through the literature synthesis. The analysis indicates that each domain represents a distinct stage in the evolution of AI autonomy while simultaneously introducing different ethical challenges for human existence. Automation primarily influences socioeconomic structures through workforce transformation, autonomous weapons raise concerns regarding legal accountability and humanitarian principles, whereas superintelligence extends ethical discussions toward long-term governance and existential risks. Collectively, these findings demonstrate that ethical concerns become progressively more complex as AI systems achieve higher levels of autonomy, reinforcing the need for integrated governance rather than isolated regulatory approaches.

### A. Ethical Findings on Automation

The reviewed literature consistently identifies automation as the most mature and widely implemented application of AI. While automation has significantly improved productivity, operational efficiency, and organizational performance across various sectors, the synthesis reveals that these technological benefits are accompanied by important ethical consequences.

One of the most frequently discussed issues concerns **workforce transformation**. AI-driven automation increasingly replaces both repetitive manual activities and routine cognitive tasks, creating concerns regarding technological unemployment, labor-market polarization, and unequal access to emerging digital opportunities (Acemoglu & Restrepo, 2020; International Labour Organization, 2023). Rather than predicting the complete replacement of human workers, recent studies emphasize that AI is more likely to reshape occupational structures by changing the skills required within future workplaces.

Another recurring theme involves **human autonomy in organizational decision-making**. As AI systems become increasingly capable of generating recommendations and predictions, organizations rely more heavily on algorithmic decision support. Although this improves efficiency, excessive dependence on automated

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

recommendations may gradually reduce human judgment, critical thinking, and individual accountability (Russell, 2019; Floridi & Cowls, 2019). Consequently, the literature consistently recommends designing AI systems that augment human capabilities while preserving meaningful human participation in decision-making processes (UNESCO, 2021).

These findings suggest that responsible automation should not be evaluated solely through productivity gains but also through its capacity to protect human dignity, promote digital inclusion, and ensure equitable workforce adaptation.

### B. Ethical Findings on Autonomous Weapons

Among the reviewed domains, autonomous weapons generate the strongest ethical consensus. The literature consistently argues that AI-enabled military technologies fundamentally challenge existing legal and ethical frameworks because autonomous systems may independently perform target identification and engagement with limited human intervention.

The most significant finding concerns the accountability gap. Existing legal principles were established for human decision-makers; however, autonomous military systems distribute responsibility among governments, military commanders, software developers, manufacturers, and AI-enabled technologies. As a result, determining legal and moral accountability becomes increasingly difficult whenever autonomous systems participate in lethal decision-making (Sharkey, 2019; International Committee of the Red Cross, 2021).

The analysis also highlights the importance of meaningful human control. Across the reviewed literature, researchers consistently argue that computational accuracy cannot replace contextual reasoning, ethical judgment, empathy, and proportionality required in military operations. Consequently, maintaining meaningful human oversight remains one of the most frequently recommended ethical safeguards for the development and deployment of autonomous weapons (Russell, 2019; UNESCO, 2021).

Overall, the synthesis demonstrates that governance of autonomous weapons should prioritize international regulation, legal accountability, and humanitarian principles to ensure that advances in military AI remain consistent with international law.

### C. Ethical Findings on Superintelligence

Compared with automation and autonomous weapons, superintelligence represents the least technologically mature but potentially the most profound ethical challenge. Although superintelligent AI has not yet been realized, the reviewed literature consistently identifies it as a strategic concern requiring proactive governance.

The dominant ethical issue emerging from the synthesis is AI alignment. Researchers argue that increasingly autonomous AI systems must remain aligned with human intentions, ethical values, and societal objectives throughout their operational lifecycle. Failure to maintain such alignment may result in unintended consequences that extend beyond technical malfunction toward broader institutional and existential risks (Russell, 2019; Hendrycks et al., 2023).

The reviewed studies further indicate that discussions surrounding superintelligence have gradually shifted from questions of technological feasibility toward governance preparedness. Rather than asking whether superintelligence will emerge, recent research focuses on how governments, researchers, and international organizations should prepare ethical frameworks capable of minimizing future risks (Bostrom, 2014; UNESCO, 2021; OECD, 2024).

Consequently, AI safety research, international cooperation, and long-term governance strategies consistently emerge as essential recommendations for ensuring that future highly autonomous AI systems remain compatible with human interests.

Table 3. Comparative Ethical Challenges Across AI Domains

| Ethical Principle | Automation | Autonomous Weapons | Superintelligence |
|---|---|---|---|
| Human Dignity | High | High | Very High |
| Human Autonomy | High | Very High | Very High |
| Accountability | Medium | Very High | Very High |
| Transparency | High | High | High |
| Fairness | High | Medium | High |
| Meaningful Human Oversight | High | Critical | Critical |

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

As presented in Table 3, the comparative analysis reveals that the intensity of ethical concerns varies across AI domains. Automation primarily raises issues related to employment transformation and organizational decision-making, whereas autonomous weapons introduce significantly greater concerns regarding accountability and meaningful human oversight because of their potential involvement in life-and-death decisions. Superintelligence exhibits the highest level of ethical concern across nearly all principles, particularly in relation to human autonomy, accountability, and long-term governance. This progression suggests that ethical governance should evolve proportionally with the increasing autonomy and decision-making capabilities of AI systems.

### D. Cross-Domain Synthesis

The cross-domain synthesis reveals that ethical concerns surrounding AI evolve in parallel with increasing levels of technological autonomy. Automation primarily transforms employment structures and organizational decision-making, autonomous weapons extend ethical issues into humanitarian law and international security, while superintelligence introduces long-term governance challenges concerning humanity's capacity to retain meaningful control over increasingly intelligent systems.

Despite these differences, the reviewed literature consistently converges on six recurring ethical principles: human dignity, human autonomy, accountability, transparency, fairness, and meaningful human oversight (Floridi & Cowls, 2019; UNESCO, 2021; OECD, 2024). These principles appear across all three ethical domains, indicating that responsible AI governance should not be developed separately for individual technologies but instead through an integrated human-centered approach.

The synthesis therefore suggests that the ethical governance of AI should be viewed as a continuum rather than as isolated technological issues. Automation, autonomous weapons, and superintelligence represent successive stages in the evolution of AI autonomy, requiring progressively stronger governance mechanisms to preserve fundamental human values while supporting technological innovation.

Table 4. Summary of Major Findings

| Major Finding | Supporting Literature | Implication |
|---|---|---|
| AI ethics extends beyond technical performance toward human-centered governance. | Floridi & Cowls (2019); UNESCO (2021) | Ethical considerations should accompany technological innovation. |
| Automation transforms workforce structures and human decision-making. | Acemoglu & Restrepo (2020); ILO (2023) | Governments and organizations should strengthen workforce adaptation and digital inclusion. |
| Autonomous weapons create accountability and humanitarian challenges. | Sharkey (2019); ICRC (2021) | International legal frameworks should preserve meaningful human control. |
| AI alignment represents the central challenge of future superintelligence. | Russell (2019); Hendrycks et al. (2023) | AI safety and international cooperation should become strategic priorities. |
| Human-centered ethical governance integrates ethical challenges across all AI domains. | OECD (2024); European Commission (2019); UNESCO (2021) | Comprehensive governance frameworks are necessary for sustainable AI development. |

Table 4 consolidates the principal findings obtained from the critical literature synthesis. Although the reviewed studies investigate different AI applications, they converge on a common conclusion that ethical governance should prioritize human-centered values rather than technological capability alone. The synthesis further demonstrates

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

that sustainable AI development depends not only on technical innovation but also on transparent governance, legal accountability, international collaboration, and the preservation of fundamental human rights. These integrated findings provide the conceptual foundation for the Human-Centered Ethical Governance Framework discussed in the subsequent section.

## DISCUSSION

### A. Ethical Transformation of Human Existence in the Era of Artificial Intelligence

The findings of this study demonstrate that ethical concerns surrounding Artificial Intelligence (AI) have evolved beyond questions of computational capability toward broader issues concerning the future of human existence. This transformation reflects a fundamental shift in the role of AI itself. Earlier generations of digital technologies primarily functioned as tools that supported human activities, whereas contemporary AI increasingly performs cognitive tasks that were previously regarded as uniquely human. Consequently, ethical debates are no longer confined to questions of technical efficiency but increasingly address the relationship between human agency, technological autonomy, and societal governance (Russell, 2019; Floridi & Cowls, 2019).

The synthesis presented in this study indicates that the ethical trajectory of AI follows a progressive pattern corresponding to increasing technological autonomy. Automation represents the initial stage of this transformation, where AI primarily assists or replaces repetitive and analytical tasks within industrial and organizational contexts. Although automation has generated significant economic benefits through increased productivity and operational efficiency, it has simultaneously altered employment structures, redefined professional competencies, and challenged traditional understandings of human labor (Acemoglu & Restrepo, 2020; International Labour Organization, 2023). These developments suggest that the ethical implications of AI cannot be evaluated solely through economic indicators but must also consider their impact on social equity, human dignity, and opportunities for meaningful work.

In addition, responsible AI governance requires organizational mechanisms capable of ensuring accountability and ethical compliance throughout the lifecycle of intelligent systems. Information governance practices, including audit trail procedures, contribute to strengthening transparency, traceability, and institutional accountability, thereby supporting trustworthy AI implementation in complex information systems (Raditya, M. R., & Nurmiati, E. 2026).

As AI systems acquire greater autonomy, ethical concerns extend beyond socioeconomic issues toward legal and humanitarian considerations. The emergence of Autonomous Weapons Systems (AWS) illustrates this transition. Unlike automation, which primarily influences labor markets, autonomous weapons directly affect decisions involving human life and death. This distinction significantly increases the ethical complexity of AI because responsibility can no longer be attributed exclusively to human actors. Instead, responsibility becomes distributed among governments, military institutions, software developers, manufacturers, and increasingly autonomous technologies (Sharkey, 2019; International Committee of the Red Cross, 2021). The literature consistently argues that meaningful human control remains indispensable because ethical judgment requires contextual reasoning, empathy, and moral responsibility that cannot currently be replicated by computational systems (UNESCO, 2021).

The progression culminates in discussions concerning superintelligence, where ethical questions shift from immediate technological impacts toward the long-term future of humanity. While superintelligent AI remains hypothetical, the reviewed literature consistently emphasizes that ethical governance should precede rather than follow technological realization (Bostrom, 2014; Hendrycks et al., 2023). Rather than asking whether superintelligence will emerge, contemporary research increasingly investigates how humanity can ensure that future AI systems remain aligned with human values. This transition demonstrates that AI ethics has expanded from reactive governance toward anticipatory governance, recognizing that preventive ethical frameworks are more effective than corrective regulation after technological deployment.

Collectively, these findings demonstrate that AI ethics has undergone a conceptual transformation. Ethical concerns no longer focus solely on individual technologies but increasingly examine how technological autonomy influences the future relationship between humans and intelligent systems. Therefore, safeguarding human existence requires governance approaches capable of evolving alongside AI itself.

### B. Ethical Implications of Automation, Autonomous Weapons, and Superintelligence

Although automation, autonomous weapons, and superintelligence differ substantially in technological maturity and application, the present study demonstrates that they share a common ethical foundation. Across all three domains, the reviewed literature consistently identifies human dignity, autonomy, accountability, transparency, fairness, and meaningful human oversight as recurring ethical principles. These findings suggest that the ethical

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY
**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

challenges associated with AI should not be interpreted as isolated technological problems but rather as interconnected dimensions of responsible AI governance.

From a socioeconomic perspective, automation illustrates how AI influences employment structures and organizational decision-making. Previous research frequently emphasizes productivity gains associated with automation; however, the present synthesis indicates that its ethical significance extends considerably further. Automation increasingly determines access to employment opportunities, influences workforce inequalities, and reshapes relationships between human expertise and algorithmic decision-making. Consequently, future labor policies should prioritize workforce adaptation, lifelong learning, and equitable access to digital technologies rather than focusing exclusively on economic efficiency (Acemoglu & Restrepo, 2020; OECD, 2024).

The ethical implications become substantially more complex within military contexts. Autonomous weapons challenge long-established legal principles because responsibility becomes increasingly difficult to assign when lethal decisions involve autonomous algorithms. Existing humanitarian law assumes identifiable human decision-makers, whereas AI-supported military operations distribute responsibility across multiple actors. This accountability gap represents one of the most significant unresolved issues identified throughout the reviewed literature (Sharkey, 2019; ICRC, 2021). Consequently, maintaining meaningful human oversight should be regarded not merely as a technical safeguard but as a fundamental ethical requirement for preserving humanitarian principles.

The discussion surrounding superintelligence further extends these concerns into long-term governance. Unlike automation and autonomous weapons, superintelligence introduces ethical questions concerning institutional preparedness, global cooperation, and humanity's future capacity to retain meaningful influence over increasingly capable AI systems. Recent studies increasingly emphasize AI alignment as the central governance challenge because future intelligent systems should consistently pursue objectives compatible with human values regardless of their computational capabilities (Russell, 2019; Hendrycks et al., 2023). These findings reinforce the importance of investing in AI safety research before technological capabilities exceed existing governance structures.

Taken together, these three domains demonstrate that ethical governance must evolve proportionally with AI autonomy. Governance mechanisms designed exclusively for current AI applications are unlikely to remain sufficient as AI systems continue to develop more advanced reasoning and decision-making capabilities. Therefore, adaptive governance capable of responding to technological evolution should become a strategic priority for governments, international organizations, industry, and academia.

### C. Toward a Human-Centered Ethical Governance Framework

One of the principal contributions of this study is the development of a Human-Centered Ethical Governance Framework integrating ethical principles across automation, autonomous weapons, and superintelligence. Unlike previous studies that typically examine these domains independently, this framework conceptualizes them as successive stages within a single ethical continuum centered on protecting human existence.

The proposed framework places human dignity at its core because every AI application ultimately affects individuals, communities, and society. Human dignity functions as the normative foundation from which other ethical principles derive. Preserving dignity requires maintaining human autonomy, ensuring fairness, promoting transparency, establishing accountability, and preserving meaningful human oversight throughout the lifecycle of AI systems.

The framework further proposes that ethical governance should operate across three complementary dimensions. The first is technological governance, emphasizing explainability, safety, robustness, and responsible system design. The second is institutional governance, focusing on regulatory mechanisms, legal accountability, and international cooperation. The third is societal governance, promoting digital literacy, public participation, workforce adaptation, and equitable access to technological benefits. Integrating these dimensions enables AI governance to respond not only to present technological challenges but also to future developments characterized by increasing autonomy and complexity.

Rather than functioning as a static regulatory model, the proposed framework should be understood as an adaptive governance approach capable of evolving alongside AI technologies. Such adaptability is particularly important because technological innovation consistently progresses more rapidly than legislative processes. Consequently, governance should remain sufficiently flexible to address emerging ethical challenges without restricting beneficial technological innovation.

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

### D. Contributions, Limitations, and Future Research

This study contributes to the growing field of AI ethics in three principal ways. First, it integrates automation, autonomous weapons, and superintelligence into a unified ethical perspective centered on human existence. Second, it synthesizes interdisciplinary literature from computer science, philosophy, law, and public policy, thereby providing a broader understanding of contemporary AI governance. Third, it proposes the Human-Centered Ethical Governance Framework as a conceptual foundation for future ethical discussions and policy development.

Nevertheless, several limitations should be acknowledged. As a Critical Literature Review, this study relies exclusively on previously published literature and therefore does not incorporate empirical evidence obtained through interviews, surveys, or case studies. Furthermore, the analysis primarily focuses on conceptual and normative discussions rather than evaluating specific AI systems or regulatory implementations within individual countries. Consequently, the proposed framework should be interpreted as a conceptual contribution rather than a validated operational model.

Future research should complement the present findings through empirical investigations examining how human-centered ethical governance can be implemented across different sectors, including healthcare, education, public administration, defense, and industry. Comparative studies evaluating national AI governance frameworks would also provide valuable insights into how ethical principles are translated into practical regulatory mechanisms. Finally, future interdisciplinary research integrating technical AI safety with ethical governance may contribute to developing more comprehensive strategies for ensuring that increasingly autonomous AI systems remain aligned with human values and societal well-being.

## CONCLUSION

This study critically examined the ethical implications of Artificial Intelligence (AI) for human existence by synthesizing contemporary literature on automation, autonomous weapons, and superintelligence through a Critical Literature Review (CLR). The findings demonstrate that ethical concerns surrounding AI have evolved alongside increasing technological autonomy. While automation primarily transforms employment structures and organizational decision-making, autonomous weapons introduce significant challenges concerning legal accountability and humanitarian principles, whereas superintelligence raises long-term questions regarding AI alignment, human autonomy, and existential risk.

The synthesis further reveals that, despite their different technological characteristics, these three domains consistently converge on six fundamental ethical principles: human dignity, human autonomy, accountability, transparency, fairness, and meaningful human oversight. These recurring principles indicate that ethical governance should not be developed independently for individual AI applications but instead through an integrated human-centered approach capable of addressing both present and future technological developments.

One of the principal contributions of this study is the proposal of a Human-Centered Ethical Governance Framework, which conceptualizes automation, autonomous weapons, and superintelligence as interconnected stages in the ethical evolution of Artificial Intelligence. Unlike previous studies that generally examine these domains separately, this framework integrates them into a unified perspective centered on preserving human existence while supporting responsible technological innovation. By emphasizing the protection of human values alongside technological advancement, the framework contributes to ongoing discussions concerning sustainable AI governance and interdisciplinary AI ethics.

The findings also provide practical implications for policymakers, AI developers, academic researchers, and international organizations. Policymakers are encouraged to develop adaptive regulatory frameworks that evolve alongside technological innovation. AI developers should integrate ethical principles throughout the AI lifecycle rather than treating ethics as a post-deployment consideration. International collaboration is equally important because the ethical challenges associated with increasingly autonomous AI transcend national boundaries and require coordinated governance mechanisms.

Nevertheless, this study has several limitations. As a conceptual Critical Literature Review, the analysis relies on previously published academic literature and international policy documents without incorporating empirical evidence from case studies, interviews, or quantitative investigations. Therefore, the proposed framework should be interpreted as a conceptual foundation that requires further empirical validation.

Future research is recommended to evaluate the implementation of human-centered ethical governance across different sectors, including healthcare, education, public administration, industry, and defense. Comparative analyses

GASET
GLOBAL ADVANCES IN SCIENCE, ENGINEERING & TECHNOLOGY

**E-ISSN: 3123-724x**
**Volume** 2, **Number** 1, June 2026
https://doi.org/10.62671/gaset.v2i1.300

of national AI governance strategies and empirical investigations into AI alignment, regulatory effectiveness, and public perceptions of autonomous technologies would further strengthen the understanding of ethical governance in the era of increasingly autonomous Artificial Intelligence.

## REFERENCES

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy, 128*(6), 2188–2244. https://doi.org/10.1086/705716

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

European Commission. (2019). *Ethics guidelines for trustworthy artificial intelligence*. Publications Office of the European Union.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1). https://doi.org/10.1162/99608f92.8cd550d1

Future of Life Institute. (2017). *Autonomous weapons: An open letter from AI and robotics researchers*. https://futureoflife.org/open-letter-autonomous-weapons/

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal, 26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines, 30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hendrycks, D., Mazeika, M., & Woodside, T. (2023). *Natural selection favors AIs over humans* (arXiv:2303.16200). arXiv. https://arxiv.org/abs/2303.16200

International Committee of the Red Cross. (2021). *Autonomous weapon systems: Technical, military, legal and humanitarian aspects*. International Committee of the Red Cross.

International Labour Organization. (2023). *Generative AI and jobs: A global analysis of potential effects on job quantity and quality*. International Labour Organization.

M. R. Raditya and E. Nurmiati, "Tata Kelola Keamanan Informasi dan Kepatuhan Etika melalui Prosedur Audit Trail,"
Journal of Informatics and Business, vol. 3, no. 4, pp. 497–504, 2026.

National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)*. U.S. Department of Commerce.

OECD. (2024). *OECD AI principles*. Organisation for Economic Co-operation and Development.

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.

Sharkey, A. (2019). Autonomous weapons systems, killer robots and human dignity. *Ethics and Information Technology, 21*(2), 75–87. https://doi.org/10.1007/s10676-018-9494-0

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Stanford Institute for Human-Centered Artificial Intelligence. (2025). *AI Index report 2025*. Stanford University.

Stix, C. (2021). Actionable principles for artificial intelligence policy: Three pathways. *Science and Engineering Ethics, 27*(5), Article 60. https://doi.org/10.1007/s11948-021-00316-7

Subiyakto, A., Yudhanta, S., **Nurmiati, E.**, et al. *Reviewing Factors of Audience Engagement in Live Streaming*. Jurnal CoreIT, Vol. 11 No. 1, 2025. DOI: 10.24014/coreit.v11i1.36039.

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.

Veale, M., & Zuiderveen Borgesius, F. J. (2021). Demystifying the draft EU Artificial Intelligence Act. *Computer Law Review International, 22*(4), 97–112. https://doi.org/10.9785/cri-2021-220402

World Economic Forum. (2024). *The future of jobs report 2024*. World Economic Forum.