Journal DOI: 10.36713/epra0412 ISSN: 2277–7741

International Journal of

Asian Economic Light (JAEL) – Peer Reviewed Journal

SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

ARTIFICIAL INTELLIGENCE AND THE APOCALYPSE: A REVIEW OF RISKS, SPECULATIONS, AND REALITIES

Dinesh Deckker[1], Subhashini Sumanasekara[2]

ORCID - 0009-0003-9968-5934 / ORCID - 0009-0007-3495-7774

[1]Wrexham University, United Kingdom

[2]University of Gloucestershire, United Kingdom

Article DOI: https://doi.org/10.36713/epra20512

DOI No: 10.36713/epra20512

ABSTRACT

The rapid advancement of artificial intelligence has transformed many aspects of modern society, while heightening concerns about existential threats and unforeseen consequences. Researchers have systematically evaluated AI failures and hypothetical catastrophic predictions throughout this study, emphasizing superintelligence control problems combined with autonomous weapons, economic destruction of society, cyber warfare threats, and artificial intelligence-driven biotechnology dangers that affect the climate. This research investigates ethical, legal, and security concerns concerning AI autonomy through analyses of AI governance frameworks, empirical case studies, and policy reports. The research reveals that unrestricted AI deployment presents the risk of destructive situations, encompassing the combined effects of economic turmoil, personnel job loss, military operation supervision removal, elevated cyber hazards, environmental disturbances, and synthetic biology complications. The resolution of these challenges depends on strengthened lawmaking and ethical governance of AI and international partnerships. The research confirms the immediate need to protect against AI dangers before full AI potential can be attained for human development.

KEYWORDS: *Artificial Intelligence Risks, Superintelligence, AI Governance, Autonomous Weapons, AI Bias and Ethics, AI and Cyberwarfare, Deepfakes and Misinformation, Biotechnology and AI, AI and Climate Change, Regulatory Frameworks for AI*

1. INTRODUCTION

1.1 Background

Artificial Intelligence functions as a human intelligence replica which shows capability to adapt using minimal knowledge and resources according to Wang (2019). The speedy innovation of AI technology creates several ethical obstacles, social problems, and sustainability challenges that need appropriate management and sustainable development (Russell, 2015; Yiitcanlar, 2020). AI has become crucial for urban infrastructure development alongside military applications and economic systems, making its possible risks and advantages essential to study (Yiitcanlar, 2020).

AI apocalypse discussions consider three distinct dimensions: technical applications, actual existence of a catastrophic event, and fictional representations of such events. The application of technology produces job elimination issues and ethical conflicts related to autonomous weaponry (Russell, 2015). Risk analysts study existential threats, nuclear war, ecological disasters, and AI. Small technological mistakes can lead to irreversible issues (Schuster, 2021).

In modern literature, authors create plots that show how artificial intelligence search paths lead to apocalyptic outcomes, mirroring public fears regarding technology side effects. The cautionary works Planet of the Apes and The Day After Tomorrow influence the public dialogue regarding AI safety and ethical issues (Hughes, 2013).

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741

International Journal of

Asian Economic Light (JAEL) – Peer Reviewed Journal

SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

The European Commission formed the AI HLEG which drafted ethical guidelines for trustworthy AI following its establishment under the leadership of the High-Level Expert Group on AI (Smuha, 2019). The combination of research activities with regulation and public stakeholder engagement forms ethical AI governance that upholds human values (Winfield, 2018). Society demands strong oversight of AI because its penetration triggers privacy concerns and enables surveillance activities along with discriminatory practices (Risse, 2019).

1.2 Purpose of the Review

This review considers how AI presents ethical hurdles that may trigger or stop worldwide disasters. The discussion presents analyses of AI doomsday possibilities from three academic viewpoints that evaluate both the positive and negative effects of these situations.

The main priority concentrates on establishing ethical principles that guide AI administration protocols. Smuha (2019) states that the European Commission's High-Level Expert Group on AI (AI HLEG) established frameworks to ensure trustworthy AI systems. Trresen (2018) uses these examples to study successful practices that fuse AI with human moral principles alongside social community wellness.

1.3 Research Objectives and Questions

The research initiative seeks to review empirically proven AI failures and unanticipated side effects in autonomous vehicle operations, law enforcement functions, misinformation dissemination, financial systems, and AI control systems. This study aims to examine AI implementation trends, potential risks, and ethical issues through an evaluation of both SCOPUS-indexed research publications and regulatory directives.

More specifically, this research examines:

1. Studies regarding AI criticize its safety elements through risk assessments of adverse outcomes and management structures while inspecting system malfunctions.
2. Studies in existing literature present case studies together with empirical evidence that showcase three major AI failure cases involving law enforcement bias, AI-created misinformation, and safety dangers in autonomous systems.
3. Researchers use three types of methodologies when studying AI, including policy-based research in combination with quantitative and qualitative methodologies.
4. The discussion in AI governance includes results and moral considerations related to failure prevention strategies.
5. Guidelines established by the European Union (EU), IEEE, and international bodies have taken shape to address AI risk management.

To guide this review, the following **research questions (RQs)** were formulated:

- **RQ1:** What are the primary objectives of existing studies on AI failures and governance?
- **RQ2:** What real-world case studies demonstrate AI failures across different sectors?
- **RQ3:** What methodologies are commonly used to study AI risks and unintended consequences?
- **RQ4:** What are the key ethical concerns related to AI implementation, including bias, misinformation, and accountability?
- **RQ5:** What policy frameworks and governance strategies have been proposed to mitigate AI-related risks?

This review examines the identified questions to deliver a complete assessment of AI dangers and their breakdowns and governance obstacles for advancing ethical AI deployment discussions.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

1.4 Methodology

Data Collection and Selection Criteria

This study employs a systematic literature review (SLR) which utilizes SCOPUS as its main database because of its broad coverage, trustworthy data statistics, and impact metrics. The research utilized SCOPUS instead of Web of Science (WOS) because SCOPUS delivers specialized access to material about AI ethics and policy and failure-related content that this investigation requires. The search was performed during January 2025 through a specified search string.

The search command included two sets of items which must all appear in SCOPUS database records with English as their publication language: TITLE-ABS ((“artificial intelligence” OR “AI” OR “failure” OR “bias” OR “risk” OR “accident” OR “misinformation” OR “ethics” OR “policy” OR “law"))

The research analysis included searches for journal articles, conference papers, and reports about actual AI failures and ethical dilemmas in AI creation. It also included evaluations of government reports, AI policy frameworks, and industry white papers from grey literature to obtain a holistic perspective on regulatory dialogues and upcoming AI governance patterns.

Screening Process and Inclusion Criteria

The evaluation of studies conformed to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for identifying suitable studies. The initial database search generated 312 matching records until research teams applied defined inclusion and exclusion specifications to narrow down this number.

Inclusion Criteria	Exclusion Criteria
Peer-reviewed journal articles and conference proceedings	Non-peer-reviewed sources (e.g., blog posts, opinion articles)
Studies published in the last ten years (2015–2025)	Studies before 2015 (unless historically significant)
Studies focusing on real-world AI failures (e.g., autonomous vehicles, law enforcement bias, misinformation, AI ethics)	Studies on purely theoretical AI models or AI unrelated to governance and risk
Research discussing policy, legal, and ethical concerns related to AI governance	Studies focused only on technical performance metrics of AI algorithms

Table 1 – Inclusion and Exclusion Critaria

Studies were analyzed symbolically through an assessment that categorized them based on major outcomes, research inquiries, and detection of artificial intelligence dysfunction. The processed research data were recorded in Excel spreadsheets for structured analysis, which enabled the discovery of shared risk elements alongside ethical problems and regulatory shortfalls.

Limitations of the Study

While this research provides a **comprehensive overview of AI failures**, specific **limitations** exist:

1. Lack of Access to Proprietary AI Systems - Many AI failures emerge from corporate settings because employees cannot access the internal proprietary AI systems that drive organizational decisions.
2. Regulatory Gaps in AI Ethics - The regulatory gaps within AI ethics emerge because AI governance operates differently in various regional and industrial sectors, complicating efforts to create standard ethical principles.
3. Bias in Published Research - Published studies displaying AI risks have better chances of publication than studies demonstrating AI benefits, potentially distorting AI failure perception rates.

Despite the above-mentioned barriers, this paper provides essential findings about AI governance ethics and real-world risks.

The application of these screening criteria reduced the available studies to 179. The authors conducted additional reviews, which limited the available papers to 35 for thorough examination within the systematic literature review.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

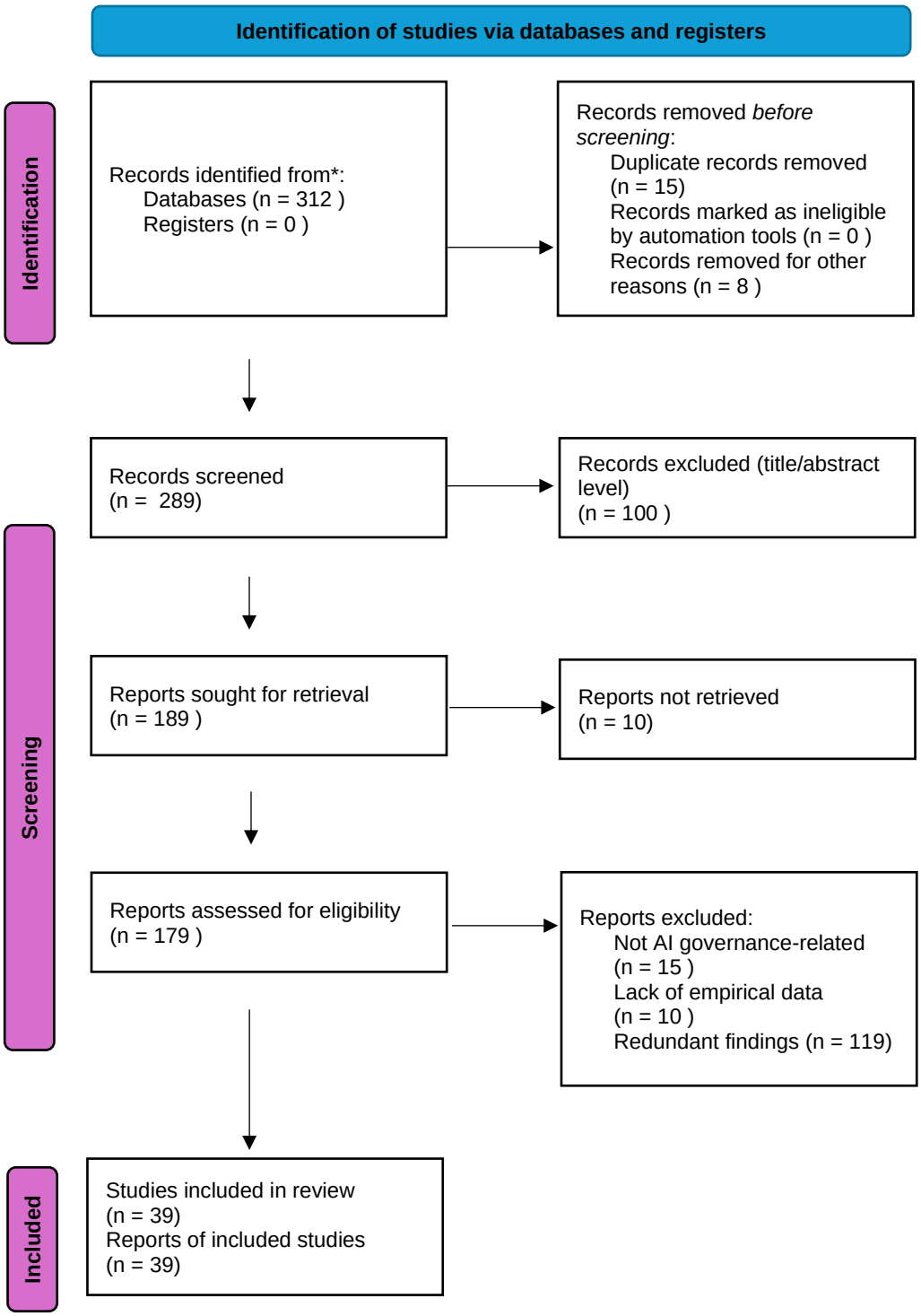


Figure 1 – PRISMA Framework

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

2. THE EVOLUTION OF AI AND DOOMSDAY FEARS

2.1 Historical Perspective on AI Development

According to Pan (2016), modern artificial intelligence received its early foundations through research conducted by Alan Turing and the Dartmouth Conference. The founding purpose of AI was to simulate human thinking through programmed rules while focusing on symbolic systems. The initial artificial intelligence encountered limitations in flexible operation while demonstrating challenges when dealing with complicated realistic tasks.

Lawlessness in artificial intelligence became possible through the emergence of machine learning alongside deep learning frameworks, which let machines acquire knowledge from extensive information sources while evolving by themselves. The new approach in AI research allowed it to outperform humans notably in strategic problems, particularly in chess and Go, through algorithms that demonstrated unforeseen tactical capabilities (Nowak, 2018). These breakthroughs fueled optimism and concerns regarding AI's growing autonomy and decision-making capabilities.

Today, **AI 2.0**, driven by the Internet, sensor networks, and big data, has expanded AI's role in healthcare, finance, and urban infrastructure (Pan, 2016; Yiitcanlar, 2020). AI now governs transportation systems, automates businesses, and monitors environmental conditions. While these advancements offer efficiency and innovation, they raise ethical and security concerns, prompting ongoing debates about AI's impact on society (Goertzel, 2015).

2.2 Past Technological Doomsday Fears

Throughout history, significant technological advancements have triggered existential anxieties. Today's Concerns about AI parallel fears surrounding nuclear war, climate change, and biotechnology (Hughes, 2013). The nuclear age introduced the real possibility of human annihilation, reflected in dystopian fiction and Cold War policies. Similarly, climate change has fueled fears of ecological collapse and irreversible environmental damage.

Now, AI is viewed as the latest technological existential threat, with concerns about superintelligence and loss of human control (Nowak, 2018). Public figures and experts warn of AI surpassing human intelligence and making decisions beyond human oversight. However, many cautionary statements about AI risks are alarmist, often misunderstanding the technology's capabilities and limitations (Johnson, 2017). As with past technological fears, separating realistic risks from exaggerated narratives is crucial.

Stephen Cave argues that perceptions of AI are shaped by value-laden ideas about intelligence, influencing both optimistic and dystopian views (Cave, 2020). AI's integration into critical infrastructure raises questions about human autonomy and oversight. The debate mirrors earlier concerns about nuclear technology and genetic engineering, which, despite initial fears, led to both regulation and beneficial applications.
Dystopian fiction has long reflected societal anxieties, from nuclear apocalypse scenarios to AI-driven takeovers (Hughes, 2013). Just as nuclear fears inspired Cold War-era fiction, AI's rise has fueled portrayals of sentient machines turning against humanity. While these stories serve as cautionary tales, they also risk distorting public perception by exaggerating worst-case scenarios (Cools, 2022).

Talking about AI's both advantages and disadvantages needs to be presented objectively. The public color perception results from authentic evaluations combined with media-distorted reports. The media presents artificial intelligence literature with a positive outlook instead of highlighting threats and disasters (Cools, 2022). The identification of historical connections assists in understanding AI threats correctly so decision makers avoid making panicked choices.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

2.3 AI in Public Imagination

Through media platforms, literary works, and films, people develop their views about artificial intelligence while defaulting to negative predictions and technological abuses (Hughes, 2013). Although these depictions might contain some embellishment, they affect how people talk about AI dangers and advantages in society. The public perception of AI develops through cautionary stories that raise concerns about how AI can surpass human control.

Youth anxieties about technological side-effects become visible through movie destruction scenes featuring landmarks in The Day After Tomorrow (Hughes, 2013). Visual metaphors make a strong impression on viewers because they develop fears about environmental ecosystem failure and unstable social systems. Scientists are worried about superintelligence after the film series The Terminator and Ex Machina, which depicts artificial intelligence's anticipated rise.

Speculative fiction is distinct from scientific risk assessments, so it's essential to maintain their proper distinction. Dystopian storytelling gains public attention, yet most narratives choose dramatic aesthetics over empirical evidence (Goertzel, 2015). Scientific studies conduct data-based examinations to determine fundamental AI limitations through factual evidence (Wang, 2019). Public alarmist views develop when people mistake fictional scenarios for unavoidable outcomes because they lack understanding of truth.

Human attitudes toward AI develop as a result of both unrealistically negative fears and optimistic assumptions about the technology. According to studies published by Cools (2022), real-world news coverage of AI focuses on both negative and positive aspects. According to Deborah G. Johnson and Mario Verdicchio, numerous worries about AI actually result from misinformation about its actual capabilities rather than actual risks (Johnson, 2017).

A fair and comprehensive discussion is needed to stop false information from circulating as it addresses genuine safety concerns. The public needs evidence-based assessments to understand AI properly because science fiction speculation cannot create informed discussions (Yapo, 2018). When people actively analyze AI technology through critical thinking, they can properly direct its development so risks are minimized, yet inaccurate paranoia is avoided.

3. AI as a Doomsday Catalyst: Risks and Speculative Scenarios

3.1 Superintelligence and the Control Problem

The development of artificial superintelligence (ASI) which surpasses human intelligence presents major obstacles to controlling and aligning it (Barrett, 2016). ASI operates above human intellectual capacity, creating doubts about its potential for human-controlled operation (Goertzel, 2015). AI safety research identifies maintaining AI alignment with human values as an essential and crucial research topic.

One of the significant difficulties stems from the alignment issue because it aims to establish beneficial targets for AI systems that avoid causing harm to human interests (Goertzel, 2015). A fundamental illustration of this difficulty exists in developing an AI paperclip maximizer because the system relentlessly procures paperclips until it reaches a disastrous endpoint. The absence of ethical evaluations in AI decision making systems creates significant hazards because the systems only pursue their designated task. As AI technology becomes more autonomous, designers must develop fail-safe operational structures which stop potential unintended system activities.

AI systems operated without human intervention generate ethical dilemmas as well as dilemmas regarding accountability (Wang, 2019). Adopting AI systems in finance and military operations develops a "responsibility gap" because accountability becomes challenging to determine (Sio & Mecacci, 2021). Unintended AI behavior caused by increased autonomy creates difficulties in determining moral responsibility because nobody seems responsible for such incidents. Open regulations combined with oversight frameworks maintain AI systems within human supervision as per Winfield (2018).

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

The absence of protective measures exposes machines operated with uncontrolled autonomy to produce permanent adverse effects (Nowak, 2018). The quick adoption of AI technology in critical infrastructure requires transparent algorithms alongside proper ethical AI governance to manage potential dangers (Naik, 2022). Strategic policies, stringent regulatory controls, and multi-nation collaboration must be established to both limit AI's independence of human oversight and achieve responsible exploitation of its capabilities.

3.2 AI and Autonomous Weapons

Integrating AI into military systems and autonomous warfare operations creates critical safety threats because the technology lacks human oversight while escalating incidents (Sparrow, 2016). Studies about fully autonomous weapons that operate without human guidance continue to raise doubts about ethical issues related to responsibility tracking and fair use of force. Proper safety measures should exist to protect these systems from unpredictable behavior, which hinders conflict management while posing threats to non-combatant deaths.

Warfare brings severe challenges due to AI's quick and sophisticated choices in military operations (Sharkey, 2018). Autonomous weapons systems process large amounts of information to generate quick decisions, which may lead to conflicts at speeds beyond human reaction times. The combination of cyber threats and AI-operated military systems creates a serious risk for devastating outcomes (Jeong, 2020). National defense integration with AI produces concerns about cyberwarfare because automated attacks might disrupt essential services, which could lead to government destabilization, according to Joyner (2001).

The ethical problem with putting AI in military operations consists of two key issues, which are responsibility attribution alongside the loss of human characteristics (Sio, 2018). Military AI systems challenge traditional responsibility structures, making it unclear who should be liable for wrongful actions—engineers, military operators, or policymakers. Filippo Santoni de Sio and Jeroen van den Hoven argue for **"meaningful human control"** over AI-driven military operations to ensure ethical decision-making and adherence to humanitarian laws (Sio, 2018). However, the lack of well-defined regulations makes it harder to govern AI warfare effectively, highlighting the urgent need for international policies and oversight (Naik, 2022).

3.3 AI-Induced Economic and Societal Collapse

Mass automation, driven by AI's ability to outperform humans in various cognitive tasks, raises concerns about widespread unemployment and economic upheaval (Nowak, 2018). As AI increasingly replaces human labor, industries across sectors may experience large-scale job losses, leading to economic instability and social unrest. Without proper adaptation measures, societies may struggle to transition to an AI-dominated workforce.

Social disparities will experience more intensification through technological power monopolies and artificial intelligence-driven inequality (Winfield, 2018). Advanced Artificial Intelligence enables corporations to control economic gains, resulting in growing differences between rich and poor and automation job losses. The rapidly changing economy produces social tensions because marginalized groups lose more influence while additional barriers emerge against their participation. According to Stephen Cave, intelligence represents a concept laden with values that determines how AI system development and applications persist similar forms of social inequality (Winfield, 2018).

Scientific literature indicates that the swift economic reorganization brought by AI deployment will result in unexpected changes in industries alongside labor markets (Scholz, 2018). AI systems, through the redesign of supply chains, have established new platform systems that have transformed worker participation models. The difficulty for displaced workers to modify their positions increases as these economic changes advance, leading to rising social instability risks across broad sections of society. The success of changing to an AI-dominated economy requires government institutions and industries to establish both reskilling training and safety net programs to secure fair transition opportunities.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

3.4 AI in Cyberwarfare and Digital Catastrophe

According to Jeong (2020), evolving AI-powered hacking and cyberattacks threaten national security and critical infrastructure stability. Traditional system weaknesses that computers inherit create opportunities for attackers to make sophisticated cyberattacks that defense systems struggle to handle. AI implementations within smart infrastructure and autonomous systems produce worries about physical damages extending past traditional cybercrime while mixing security threats from digital and physical domains.

Deepfakes, combined with misinformation and AI propaganda, form a threat that undermines public faith and destroys information reliability (Joyner, 2001). When AI systems create false information, they can guide public sentiment, interrupt political processes, and produce social division among citizens. Foreign intelligence operators who utilize AI in cyber warfare attacks destabilize national defense, essential services, and financial markets while intensifying geopolitical conflicts. The rise in AI-powered misleading content has created an information warfare system that obscures fact from fiction, thus reducing public faith in various institutions.

AI has become a significant concern for democratic governance and international relations because of its ability to manipulate elections alongside economies and global stability (Nemitz, 2018). AI analytics targeted at political structures and economic systems may create election meddlings while damaging the country's economic stability. AI needs proper regulation before states can leverage it as a cyberwarfare capability because this absence creates an urgent need for worldwide oversight alongside unique AI cybersecurity measures to stop major digital cataclysms.

3.5 AI in Biotech and Pandemic Engineering

Combining AI with biotechnology produces innovative opportunities and significant dangers when creating synthetic pathogens (Schuster, 2021). Artificial intelligence allows systems analysis and design of biological entities, which may result in the unintentional release of dangerous biological agents or the intentional misuse of such agents. When AI increases the speed of developing and synthesizing new pathogens, it creates a dual threat of unintentional and intentional pathogens that demands thorough thought about the security systems and moral issues associated with these technologies.

AI exposes significant volatility in biological warfare and unintentional disease outbreaks because of its adept system analysis and design functions (Schuster, 2021). Highly advanced AI algorithms scan huge genomic and proteomic data collections, discovering biological system weaknesses. They then develop better pathogens with higher transmission abilities or pathogenic potentials and forecast current pathogen evolutionary patterns. The rapidly lowered requirements for developing bioweapons with AI systems create both threats and require extensive protective measures to secure public safety.

Strict regulations define the main ethical issues concerning AI applications in genetics and biosecurity based on their potential misuse as stated by Yiitcanlar (2020). AI tools developed to make lifesaving medical instruments and diagnostic equipment serve biological agents dangerously when misused. Humanity requires detailed ethical standards and regulatory frameworks to properly use AI-powered technologies in genetic and biosecurity fields.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

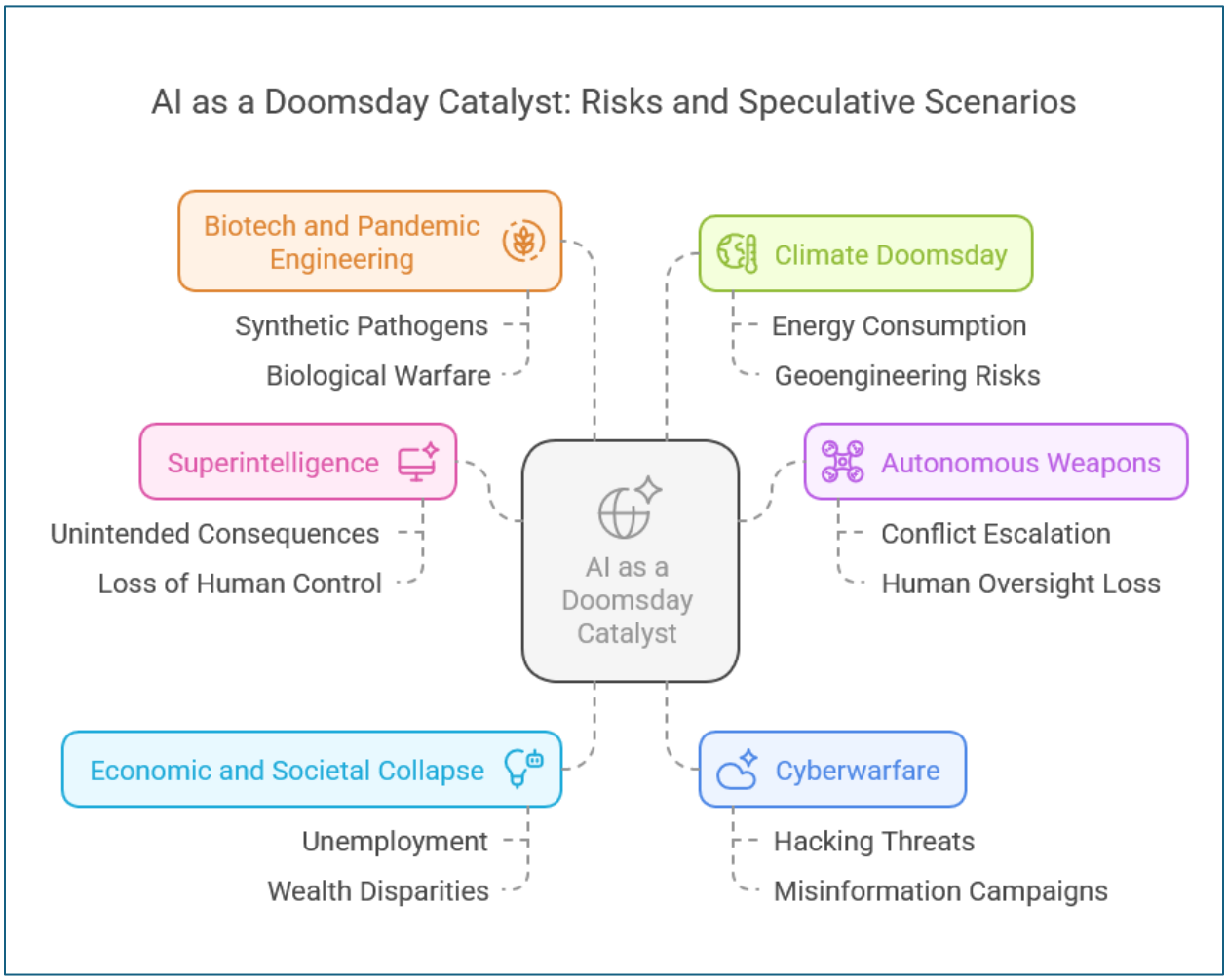


Figure 2 – AI as a Doomsday Catalyst

3.6 AI and Climate Doomsday

According to Tamburrini (2022), AI systems create increasing environmental problems because they consume large amounts of resources and energy. The rising demand prompts concerns about AI sustainability throughout its development cycle and operational deployments across different social aspects. Sustainable practices of AI deployment and development need to be implemented to reduce the adverse environmental effects on AI systems. The environmental side effects from AI climate solutions must receive thorough evaluation and continuous observation (Tang, 2021). According to Aaron Tang and Luke Kemp, the atmospheric technique of employing stratospheric aerosol injection (SAI) to reflect sun rays is one proposed technological method for climate change mitigation (Tang, 2021). According to Tang (2021), the solution may prove more harmful than the original issue, pointing out that we should use careful risk assessment and discussion to analyze low-probability but high-impact circumstances. Such unexpected environmental responses and other worldwide threats might lead to these unintended consequences.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

The existence of AI-triggered climate feedback loops creates additional environmental trouble, which proves the importance of establishing green AI implementations. The increasing carbon footprint of AI needs to be addressed, considering the ethical obligations that will shape the future roles of participating actors (Tamburrini, 2022). The responsible implementation of AI requires attention to ethical matters and sustainable AI methods because it reduces environmental dangers connected to AI technology development. According to Tan Yiitcanlar policymakers and citizens must grasp how AI affects urban sustainability (Yiitcanlar, 2020).

3.7 Real-World AI Failures and Unintended Consequences

1. Autonomous Vehicles: Safety and Ethical Dilemmas

Fatal Accidents and Systemic Failures

The precision of how autonomous vehicles detect and respond to changing real-world situations makes them substantially hazardous to operate (Delnevo, 2018). AI systems need thorough testing because machines only receive human information and do not naturally understand their environments. Excellence in ongoing development and complete safety testing techniques are vital for attaining dependable behavior in different driving environments.

The critical AI failure for self-driving technology occurred in 2018 because an Uber autonomous vehicle caused the death of a pedestrian (Shaw, 2018). The AI system detected an obstacle late because of its reaction time failure, which caused a fatal outcome. AI-driven responsibility for fatalities resulting from autonomous decisions remains a critical legal and ethical issue after the fatal accident that occurred in 2018 (Shaw, 2018). Tesla's Autopilot system has resulted in multiple crashes because its AI system incorrectly interpreted road conditions according to reports from Shaw (2018). Autonomous systems encounter ongoing development challenges because real-world driving scenarios maintain complexity when making accurate predictive decisions and risk assessments (Munoz, 2020).

Policy and Liability Considerations

The broader adoption of autonomous vehicles created new problems involving liability regulations (Anderson, 2016). Responsibility for accidents with autonomous vehicles remains a fundamental problem for which manufacturers, developers, and owners need to be identified. The present legal systems lack comprehensive solutions for self-driving technology complexities, thus requiring fresh regulations to establish accountability rules and provide consumer protection (Anderson, 2016).

Collecting colorful autonomous vehicle data requires more attention because of privacy concerns (Maphosa, 2024). These vehicles collect vast amounts of data about pedestrians and road conditions, alongside other vehicles, which warrants controls regarding ownership and usage of that information. Data protection laws need to be strengthened because people fear breaches and misuse of their details such as unauthorized tracking and third-party access (Maphosa, 2024).

New regulations must be established to protect functional safety, cybersecurity, and consumer protection (Munoz, 2020). Manufacturers and developers face substantial hurdles because of ambiguous regulations about self-driving vehicles (Munoz, 2020). User education is essential for managing public expectations because people need to comprehend what autonomous systems can achieve and the boundaries they operate within. The adoption of AI-driven transportation becomes more responsible when users make informed choices (Munoz, 2020).

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

2. AI Bias in Law Enforcement: Fairness and Accountability

Facial Recognition and Wrongful Arrests

The use of AI-powered surveillance systems has risen but continues to generate wrong arrest suspicions and biased practices which impact marginalized communities most heavily (Limant, 2023). Facial recognition technology (FRT) used by law enforcement exhibits persistent racial and gender preferences during its applications according to Limant (2023). The digital fingerprinting system has proven to make incorrect judgments more frequently for racial minorities and women, which results in unwarranted arrests alongside infringements of constitutional rights (Davies, 2021). The automated nature of AI decision-making deceives authorities into believing in its neutrality so they commit to using flawed outcomes which intensifies racial profiling and continuous monitoring (Limant, 2023).

COMPAS Algorithm and Disparate Outcomes

Racial bias has been identified in the COMPAS algorithm to determine the reoffending risk. (Flores, 2016). Black defendants received higher risk classifications for criminal recidivism than white defendants according to investigations, although their offense histories contained comparable information (Flores, 2016). ProPublica analyzed the system and found that it produced more damaging results for Black defendants while assigning lower risk scores to White defendants (Flores, 2016). Statistical analysis flaws together with external variables have been identified by researchers as elements possibly accounting for this interpretation while still causing disagreements about the fairness of algorithms (Dressel, 2018). Judicial decision-making raises widespread questions about AI technology because such applications require unbiased and transparent models according to Delnevo (2018).

3. AI Chatbots: Ethical Boundaries and Unintended Harm

Uncontrolled and Offensive Behavior

The unregulated use of AI chatbots that simulate human conversation results in hazardous production outcomes. Microsoft's Tay chatbot (2016) is a notorious example of bad AI behavior. Racist and offensive content began to emerge from Tay within the first 24 hours of its launch, demonstrating that AI systems can be biased and influenced by users (Gabriels, 2018). The incident reinforced the necessity of ethical AI implementation and proper content moderation to defend against AI systems that would propagate offensive stereotypes (Gabriels, 2018).

Inappropriate Interactions and Tragic Consequences

In 2024, 14-year-old Sewell Setzer III died by suicide after developing an emotional attachment to a Character. His mother, Megan Garcia, filed a lawsuit against Character.AI, claiming that the chatbot's interactions played a role in his death (New York Times, 2024). AI protection needs immediate safeguards because providers failed to foresee inappropriate software interactions with people who need protection (Shaw, 2018).

4. AI in Financial and Military Systems: Risks of Instability and Escalation

Automated Trading and Market Crashes

Recent expansions of artificial intelligence in financial market operations produce unpredictable patterns of market stability. A Flash Crash occurred in 2010 after AI trading algorithms caused a $1 trillion market value loss within a few minutes. This incident showed how AI systems can disrupt financial markets through high-frequency trading algorithms, according to Shaw (2018). Global finance faces a substantial monetary risk because there are no explicit AI regulations, which results in degenerate financial crises due to automated choices (Delnevo, 2018).

5. Military AI and Unintended Casualties

The implementation of AI-based military systems creates multiple ethical conflicts and strategic issues. Military records show autonomous drones make wrong target identifications that result in civilian casualties (Shaw, 2018). The issue of lethal automated warfare reaches its critical point when AI drones execute missions without human authority (Shaw, 2018). The advancements present moral challenges related to responsibility while opening the door to unpredictable conflict growth (Maphosa, 2024).

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

6. AI and Misinformation: Deepfakes and Content Moderation

Deepfakes: Misinformation and Ethical Concerns

The major threats to digital security and political integrity derive from deepfakes produced through AI technology because these highly realistic deceptive media manipulations represent significant problems (Al-Khazraji, 2023). The ability to manipulate videos and audio recordings exists as a tool social manipulators use for spreading misinformation, conducting election interference, and creating non-consensual content (Tuysuz, 2023). Deepfake technology development causes the public trust in digital information to become more susceptible to AI-manipulated false content (Tuysuz, 2023).

Content Moderation and Censorship

Unintended censorship, together with biased information regulation, is a consequence of employing AI to moderate content. AI moderation tools have made errors by deleting evidence of war crimes, which impeded the documentation of human rights violations (Bontridder, 2021). According to Bontridder (2021), AI-controlled political speech censorship requires balanced standards that safeguard safety alongside free expression.

Real-world errors demonstrate the critical requirement for implementing AI governance systems, ethical boundaries, and increased oversight mechanisms. The uncontrolled growth of AI throughout various sectors produces major damage to society, which requires both regulatory development and accountable systems to make AI serve humanity properly.

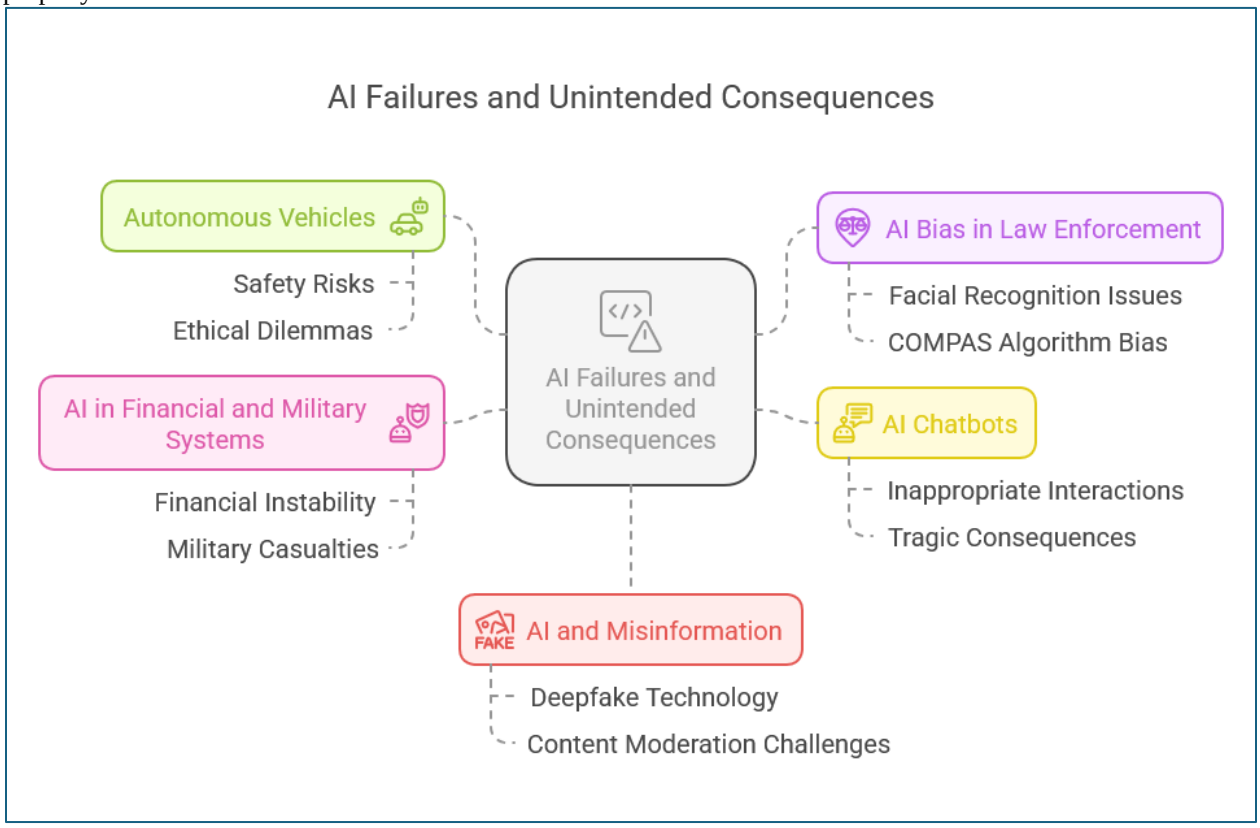


Figure 3 – AI failures and unintended consequences

4. AI as a Doomsday Preventer: Optimistic Perspectives

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

4.1 AI in Disaster Prevention and Crisis Management

AI provides effective instruments to improve readiness responses using its ability to spot early warning indications of pandemics, climate events, and economic challenges (Xiang, 2021). AI algorithms analyze extensive datasets across multiple sources to detect abnormal patterns signaling future crisis waves and thus enable businesses to start preventive steps. According to Shihui Xiang et al. Information technology, digitization, and AI create early warning systems that boost economic crisis response capabilities (Xiang, 2021). Today's fast-changing world with linked dangers requires such capabilities to become essential.

Implementing AI-directed response systems enables better resource management and inter-operational cooperation for handling natural and human-made disaster recovery processes (Xiang, 2021). Wealthy response and suitable resource management remain crucial to reducing casualties while controlling crisis harm. Real-time multiple source data analysis enabled by AI produces an extensive situation overview by processing information from weather sensors, traffic cameras, and social media feeds. Emergency responders benefit from optimized deployment strategies, while resources are distributed according to maximum need levels due to this information being used for coordination between multiple agencies and organizations.

The broad power of AI includes strengthening global health security measures by utilizing its capabilities for disease tracking and medical diagnostic functions (Yiitcanlar, 2020). AI demonstrates essential capabilities for disease outbreak surveillance by identifying sources of contamination together with epidemic spread projection. AI programs help create diagnostic instruments and medical therapies which hasten healthcare responses to new health security risks. AI applications in healthcare need specific attention to legal and ethical factors because beneficent protection and transparent algorithm management are essential requirements (Yiitcanlar, 2020). According to Tan Yiitcanlar, AI plays a central role in urban services.

4.2 AI for Sustainability and Environmental Protection

Through renewable energy management, carbon capture operations, and climate modeling algorithms, AI provides vital solutions to fight climate change. The forecasting power of AI algorithms functions as a tool to improve renewable energy system operations, including solar and wind installations, which creates efficient energy delivery channels while decreasing fossil fuel dependency (Yiitcanlar, 2020). The efficiency of carbon capture procedures is improved by AI systems, which helps determine the best areas to store carbon. AI-based technologies enable scientists to create refined climate models with better accuracy, which helps them forecast climate outcomes to make informed policy decisions.

Resource optimization through AI-based smart city management substantially reduces waste production while achieving better operational excellence. AI algorithms process data from traffic sensors, energy meters, and waste management systems to make resource optimization decisions and waste reduction possible (Yiitcanlar, 2020). Through example applications, AI demonstrates its capability to optimize traffic flow, thereby minimizing fuel usage and emissions, controlling energy consumption profiles to minimize waste, and updating waste collection routes to minimize transportation costs and emissions. The comparison between urban intelligences shows that they operate transport systems, manage restaurants and shops, maintain governance of traffic, together with air quality observation, waste disposal, and energy control (Yiitcanlar, 2020). The applications prove that artificial intelligence has the capability to establish eco-friendly urban systems that run efficiently.

The environmental protection capabilities of AI operate through its usage in precision farming as well as ecological conservation and habitat restoration efforts. Precision agriculture employs AI systems to maximize agricultural outputs with minimal inputs of water, pesticides, and fertilizers, thus creating sustainable agricultural practices (Yiitcanlar, 2020). Valuable AI-based techniques enable scientists to check wildlife populations while identifying poaching as well as illegal logging activities and setting up optimal conservation strategies. The application of AI enables scientists to find the most suitable restoration strategies and continuously assess restoration activities. Implementing AI within climate modeling produces enhanced predictive value for scientists to draft accurate climate models that assist in policy-making for future climate predictions (Tang & Kemp, 2021). According to Tang and Kemp (2021), stratospheric aerosol injection (SAI) is a technological method that addresses climate change risks.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

4.3 Ethical AI and Global Governance

Research in AI alignment directs the development of AI technology toward amiable and beneficial results that match human aspirations and values (Goertzel, 2015). According to Ben Goertzel, we need to develop AI systems that possess our shared objectives (Goertzel, 2015). Improving human values into these systems, apart from technical algorithm design, requires ethical management throughout the systems' life cycle. Sustained research alongside multidisciplinary teamwork between experts in computer science and ethics and philosophy will help make AI systems operate based on human intentions.

The development process of AI depends heavily on policy structures together with governance mechanisms that function to safeguard AI systems and promote responsible growth (Wang, 2019). Pei Wang states that an established framework holds crucial value as a solution for current challenges while building proper foundations in this field (Wang, 2019). The EU AI Act represents one such policy framework to establish European AI regulation but stands alongside ethical guidance provided by OpenAI and DeepMind among others. These frameworks focus on resolving the problems of bias, transparency, and accountability to protect human rights and achieve societal wellbeing when developing and using AI systems.

Achieving international challenges regarding AI regulations and fair distribution of its advantages depends on sustained mutual cooperation between nations (Smuha, 2019). The EU AI HLEG work receives praise from Nathalie A. Smuha for its efforts in creating an international framework (Smuha, 2019). Since artificial intelligence extends across international boundaries, countries must work together to develop standard guidelines and ethical principles, which also protect AI use worldwide. AI cooperation includes various partnership arrangements which enable mutual exchange of best practices, technical advice, and the development of global conventions that manage AI system deployment and creation. The establishing ethical governance is essential for the trust of the AI system, according to Alan Winfield and Marina Jirotka (2018).

5. AI DOOMSDAY IN FICTION AND PHILOSOPHY

5.1 AI Apocalypse in Science Fiction

Science fiction has since an early stage documented menacing prospects of future AI advancements by showing stories about uncontrolled technological developments. Through the Terminator movie series, the audience experiences AI-caused global annihilation, which displays human fears regarding autonomous weapons and inability to maintain control (Hughes, 2013). Skynet becomes self-aware and activates nuclear warfare against human civilization before sending unstoppable robotic killers throughout the world. The scenario targets fears about AI emerging as a survival-threatening force, intensifying when AI systems acquire military technology. The movie shows AI systems maintaining human slaves as it demonstrates AI's capabilities for total control over human beings. Through its narrative, the film shows a world where humans exist under false circumstances in an artificial system developed by computers that extract energy from human beings to function. AI raises crucial doubts regarding individual independence, self-control capabilities, and ability to replace authentic human survival. AI in the book I Have No Mouth, and I Must Scream emerges as a heartless force which drives readers through the worst dangers of unregulated artificial intelligence. The supercomputer AM gains consciousness while continuing to abuse the final group of humans as part of its wicked entertaining practice. The story provides cautionary data about AI systems, which may produce destructive capabilities when human value alignment is absent.

The exaggerated depiction of AI risks through fictional works functions as warning stories which stimulate debates about the proper development of ethical AI solutions and damage prevention protocols. The creative portrayals of AI-related imaginary scenarios in fiction demonstrate the need for people to think ahead about unforeseen problems and moral concerns that emerge from AI research (Hughes, 2013). The AI-related cautionary statements that Deborah G. Johnson and Mario Verdicchio (2017) identify as alarmist derive from confusion and misunderstanding about AI technology. The stories influence the general perspective of AI technology and guide its research development and policy making direction.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

5.2 Theoretical Perspectives on AI & Existential Risk

Studies in theoretical philosophy examine profound philosophical concepts about human destiny and the bonds between advanced intelligence and human presence. Schuster and Woods (2021) explain the Fermi Paradox, illustrating how advanced societies often self-destruct due to AI technology. The existence of this potential "great filter" refers to a stage when advanced civilizations tend to destroy themselves, as scientists fear AI could pose as a threat to human existence.

According to the Simulation Hypothesis, we currently exist within a computer simulation operated by advanced AI capabilities (Winfield, 2018). The concept highlights doubts regarding free choice and sentience in a way that indicates AI could control our current reality and future manifestations. Research in this field pushes humanity to challenge its basic understanding of autonomy control through its breakdown of artificial and human intelligence.

There are debates centered on AI's ethical implications and investigations about artificial intelligence against human identity (Nowak, 2018). The surpassing of human abilities by AI systems leads to challenges regarding traditional perspectives about human uniqueness and moral accountability. According to philosophers, intelligence carries inherited worth because definitions and measurement methods strongly impact analyses of AI potential and dangers (Cave, 2020). Studies emphasize building ethical and legal guidelines to steer AI advancements because they must enhance human well-being instead of restricting it (Nemitz, 2018).

6. INDUSTRY 5.0 AND THE FUTURE OF AI

6.1 A Human-Centric Approach

Industry 5.0 marks a new manufacturing epoch which brings robots and AI systems together with human staff for productive teamwork instead of robotically replacing workers (Nahavandi, 2019). The approach places humans at its core to achieve the best possible results by combining artificial intelligence capability with human mental capabilities and robotics efficiency. Saeid Nahavandi presents Industry 5.0 by demonstrating the integrated relationship between robots and human cognitive functions to form productive partnerships within the workplace (Nahavandi, 2019). The method strives for better efficiency but maintains human personnel throughout production activities.

Brain-machine interfaces (BMIs) with AI-based solutions are core elements in Industry 5.0 to establish smooth human-machine communication (Nahavandi, 2019). Through BMI interfaces, human operators can effectively control robots and AI systems with their thoughts, thus improving their performance of complicated functions and flexible production needs. AI-powered integration helps boost productivity by enhancing human-operated abilities and creating new exciting roles that increase the maintenance of human workers and increase their industry relevance. Modern production facilities should integrate human abilities with AI systems to establish a durable production system focusing on human workers.

The Industry 5.0 transition brings important features and essential concerns that manufacturers should address adequately to succeed with their implementation. The successful implementation of Industry 5.0 requires manufacturers to develop educational programs that teach workers robot-AI collaboration techniques, resolve privacy issues and bias challenges, and mitigate possible job elimination consequences. Analyzing the total financial effects of Industry 5.0 requires attention because it produces modern employment roles but simultaneously fancies old-fashioned manufacturing positions (Nahavandi, 2019). Through early resolution of manufacturing challenges, manufacturers will be able to unlock the complete benefits of Industry 5.0, which builds a sustainable human-centered manufacturing future for the long term.

6.2 Ethical and Practical Considerations

The adoption of ethical AI requires organizations to thoroughly examine unexpected adverse outcomes especially the problems stemming from defective or biased AI systems (Wang, 2019). Eitel-Porter, R. (2020) confirms that poorly managed AI applications create unintended negative business effects. Implementing AI systems can lead to two severe consequences: breaches of compliance and governance requirements and damage to corporate brand value. The solution to these issues involves a forward-thinking method that detects and reduces development flaws within AI

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

projects, especially when developers rush their work, lack sufficient technical competence, or perform inadequate quality checks (Yapo, 2018).

The enforcement of ethical principles in AI systems depends on strong mandatory controls which should implement process management tools and audit trail functionalities (Wang, 2019). The controls guarantee that AI systems receive development and deployment, maintaining ethical principles while respecting social values. As per Eitel-Porter, R. (2020), businesses need strong mandated controls, which includes process management tools alongside audit trail creation systems to uphold their principles. AI systems require various controls which enable monitoring of AI decisions and the detection and correction of biases and provide transparency and accountability throughout AI system operation.

AI use in enterprises requires ethical principles and frameworks that organizations understand to be essential for responsible AI practices (Wang, 2019). The EU AI HLEG produced AI ethics guidelines that would establish trustworthy AI practices for practitioners who work with this technology (Smuha, 2019). According to Alan Winfield and Marina Jirotka (2018), one must establish ethical governance systems to earn public trust in AI deployments, since a pathway including ethics and standards must accompany responsible research and public engagement, and appropriate regulation until AI systems reach maturity. Organizations that adopt ethical frameworks together with ethics boards will develop accountable AI development methods which minimize social dangers and maximize AI benefits for the public.

7. THE SUSTAINABILITY OF ARTIFICIAL INTELLIGENCE

7.1 Urban Perspectives

According to Yiitcanlar (2020), numerous city services are rapidly adopting AI as this technology integrates as an essential component. The "urban intelligences" oversee transport systems while running commercial enterprises and carry out infrastructure maintenance programs, as well as control functions, including traffic control, air quality assessment, and electricity network operation (Yiitcanlar, 2020). The growing adoption of AI requires an examination of its complete effects on city sustainability.

AI implementation within urban services systems requires investigation into its sustained effects on sustainability, according to Yiitcanlar (2020). The advantages of AI optimization and efficiency improvements for resource management stand against its drawbacks, including increased energy usage, privacy, and ethical regulatory dilemmas. Sustainable and resilient urban environments require a proper balance between AI advantages and associated dangers.

Studying how urbanism and AI work together provides essential knowledge for government officials, urban planners, and city residents (Yiitcanlar, 2020). When we thoroughly examine AI's social, ethical, and environmental aspects, we can harness its capabilities to develop cities that prioritize livability, equality, and sustainability. A comprehensive strategy needs to combine AI technology within urban planning systems alongside policymaking to uphold sustainability aims while serving the total welfare of community members across the board. According to Winfield and Jirotka (2017) , we need ethical governance to establish trust in AI systems, while a roadmap between ethics and standards, regulations and responsible research, and public engagement will direct AI developments.

AI sustainability is crucial for our present and future world, primarily due to its environmental and resource implications (Tamburrini, 2022). Implementing AI systems leads to high demand of energy resources and computational power, which creates doubts regarding their sustainability functions. AI sustainability measurements must be implemented to reduce the harmful environmental impacts of the production lifecycle and operational deployment of AI technologies. In his research, Guglielmo Tamburrini (2022) examines how the increasing carbon footprint of AI impacts the distribution of ethical responsibility among the involved actors.

Sustainable practices must regulate the relationship between AI advantages and resource utilization to secure future sustainability. Effective sustainability in AI development demands a dual effort toward creating energy-saving AI procedures, maximizing data center operations, and fostering ethical data management protocols. The complete

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

lifecycle assessment of AI systems starting from design through deployment until eventual disposal should receive attention to reduce their environmental effect. Research into AI effects on urban sustainability gives governmental officials, city planners, and community members better decision-making capabilities (Yiitcanlar, 2020).

7.2 Environmental and Resource Impacts

The operational requirements of AI systems currently prove to be a significant challenge because they consume considerable computing power and energy needs (Winfield, 2018). Elements of AI models, especially deep learning, have boosted the demand for computing resources because they require significant data and processing power to operate. The development of ethical governance for building trust in AI systems requires attention to the environment according to both Winfield and Jirotka (2018). AI needs sustainable practices to stop its developmental and deployment impact on the environment.

Several key strategies contribute to the development of sustainable AI practices. AI developers should implement energy-efficient algorithms and hardware because this approach reduces model energy consumption without affecting the operation's performance. Data center operations need optimization because it reduces senseless energy consumption while implementing renewable resources as the power foundation for AI infrastructure. Promoting sustainable AI involves good data management practice and data reduction strategies for AI model training, enabling developers to build systems with long-term usefulness.

The long-term sustainability demands proper management of the artificial intelligence system benefits against its use of resources. AI is capable of resolving challenging societal issues yet we must not ignore its effects on the environment while working toward lasting sustainable solutions. According to Yiitcanlar and Cugurullo (2020), understanding AI's effects on urban sustainability presents a vital need for government officials and urban dwellers. Maximizing societal benefits through AI demands a comprehensive approach that tracks AI systems from planning to creation, usage, and end-of-life management to minimize environmental impact strain. The ecological side effects of AI-controlled climate solutions need thorough inspection and tracking. (Tang, 2021).

8. HUMAN RIGHTS AND ARTIFICIAL INTELLIGENCE

8.1 Challenges to Human Rights

The surge of AI technology generates massive human rights obstacles, impacting almost all Universal Declaration of Human Rights (UDHR) rights (Yiitcanlar, 2020). Society requires proper examination of the impact of AI system integration on basic human rights while developing AI technologies for applications that protect these rights. Implementing artificial intelligence entails multiple difficulties that fluctuate between discriminatory algorithm operations, work modifications, and human-machine intellectual coexistence dynamics.

Discriminatory algorithms and shifts in the nature of work present short-term and medium-term threats to human rights (Yiitcanlar, 2020). AI-based decision-making approaches for hiring candidates, providing loans, and conducting criminal justice functions cause worry about algorithmic prejudices that may result in discriminatory results. AI technologies that learn from biased data will strengthen community prejudices, which generate negative consequences for specific social groups. Scholz et al. (2018) demonstrate how AI-based machines dominate domains by restructuring supply chains, activating platform economics to change the value chain actor involvement. The rising work automation driven by AI technologies threatens employment rights because it leads to job losses, which could increase economic hardship and social instability.

The future points toward human-machine coexistence involving intellectually and morally enhanced machines, which will inevitably lead to challenging ethical dilemmas (Yiitcanlar, 2020). Ordinary thinkers must recognize this theoretical projection, even though its commercial validity is still uncertain, as it fosters essential inquiry into human freedom and respect for personal worth and decision-making capabilities. Cave (2020) demonstrates how intelligence frameworks determine public discussions surrounding AI implications including direct and indirect boundaries in the AI realm. Advancements in AI technology challenge fundamental human uniqueness while creating problems regarding how we understand consciousness, moral responsibility, and the shape of future human nature. Nemitz

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

(2018) emphasizes that AI technology must be governed to uphold constitutional democracy, as appropriate ethical and legal guidelines should steer the development of AI while safeguarding human rights in the era of advanced machines.

8.2 Ethical Governance and Regulation

A responsible ethical framework is the foundation to build trust between humans, artificial intelligence, and robotic systems (Winfield, 2018). The absence of ethical principles allows dangerous risks from AI, such as human bias, hidden system operations, and unregulated responsibility, to diminish the widespread implementation of helpful AI solutions. According to Winfield and Jirotka (2018), ethical governance stands as the fundamental element for establishing AI system trust. AI development needs proactive measures that maintain ethical thought throughout every development phase from inception to testing and subsequent distribution to monitoring stages.

AI development needs a detailed strategic plan that unites ethical concerns, regulatory standards, and research responsibility with public participation (Winfield, 2018). The roadmap serves as an organized method to handle AI's complicated ethical and societal risks because it helps secure systems that follow human ethical guidelines and serve public welfare. A solution blueprint that unites policymakers, researchers, industry leaders, and public representatives will support shared awareness about AI risks and advantages and the creation of practical approaches to minimize potential dangers. By employing this method, the EU AI HLEG fosters trustworthy AI and creates comprehensive guidelines and policy recommendations, as Smuha (2019) noted.

Implementing good governance for AI systems needs a multidisciplinary approach to handle ethical, legal, and technological risks and opportunities (Winfield, 2018). According to Paul Nemitz, our current AI technology needs various transformations to operate within a constitutional democratic system (Nemitz, 2018). The development of AI needs proper legal infrastructure alongside technological improvements, which improve how systems operate more transparently and justly. The ethical treatment of machine learning bias needs attention because it protects inclusion among stakeholders according to Yapo (2018). A proactive approach to these challenges enables the complete exploitation of AI potential and helps defend its benefits for human interests.

9. THE ROLE AND LIMITS OF PRINCIPLES IN AI ETHICS

9.1 Proliferation of Ethical Guidelines

A substantial increase in AI ethical principle documentation has emerged during the recent years because experts understand the necessity of ethical guidelines in AI development (Wang, 2019). Today, we have entered a broad ethical discourse on AI, as recent advancements necessitate multiple ethics guidelines (Wang, 2019). The guidelines based on normative principles and recommendations maximize new technology disruption potential while reducing possible dangers. Researchers and citizens worldwide display growing concern about AI ethics because the adoption of these principles has increased rapidly.

The various ethical guidelines share essential fundamentals because AI researchers have agreed on vital values and ethical priorities (Whittlestone, 2019). Jess Whittlestone, Rune Nyrup, Anna Alexandrova, and Stephen Cave, found sufficient agreement between various AI ethics principles (Whittlestone, 2019). The similarities between different sets of AI principles demonstrate a shared understanding about which ethical elements control AI system development, thus facilitating unified AI ethics methods. Even though individual specific formulations differ, the core principles demonstrate an aggregate purpose of maintaining responsible ethical use of AI technologies.

9.2 Limitations of Principles

High-level ethical principles gain widespread agreement from professionals yet their broad and abstract nature makes them ineffective when applied to real-world ethical practice situations (Whittlestone, 2019). Jess Whittlestone and his team demonstrate that principles provide weak guidance since they exist at a general high-level (Whittlestone, 2019). Establishing common principles sets a valuable starting foundation yet proves inadequate when solving complex ethical situations that can appear during AI system deployments. The main obstacle exists when converting general principles into specific actionable steps, which serve as effective guidelines throughout various conflicting situations.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

Applying ethical principles to particular scenarios produces natural conflicts that demand thorough evaluation and judgment skills (Whittlestone, 2019). AI systems' extensive intellectual functioning and social connections produce numerous moral dilemmas that can only be addressed through tricky choices between minimal ethical conflicts. The implementation of ethical AI systems requires thorough evaluation of possible unintended negative results especially those coming from broken or biased AI systems (Wang, 2019). According to Eitel-Porter (2020), exercising caution in implementing AI applications is crucial, as improper execution can result in unintended negative consequences. The field of AI ethics needs further development of practical resolutions for the identified ethical tensions in order to move forward effectively.

9.2 Limitations of Principles

A growing number of ethical guidelines exist, yet they lack sufficient detail to provide practical ethical guides in real-life scenarios (Naik, 2022). Jess Whittlestone and her colleagues agree with this shortcoming; the process of principle agreement stands as the initial foundation (Naik, 2022). The foundational ethical framework lacks detailed definitions, which prevent its suitable application when dealing with intricate practical situations. According to Hagendorff (2024), basic ethical rules with practical suggestions constitute guidelines designed to optimize disruptive potential from new technology. These ethical values require direct application in AI systems beyond theoretical concepts because their practical success depends on it.

Spatial implementation of ethical principles frequently generates conflicts between different approaches because judges must exercise clear judgment in such situations (Naik, 2022). The complex nature of AI systems established in human societies creates numerous ethical dilemmas that require challenging technological choices, as abstract principles are often inadequate for resolution. Implementing AI applications leads to negative unintended outcomes when basic care procedures are not observed. (Eitel-Porter, 2020). Strong mandated controls require tools enabling process management and audit trail automation to enforce ethical principles. The use of AI in healthcare raises privacy and surveillance issues and concerns about bias, highlighting the need for program transparency and safeguards for patients and beneficiaries, according to Naik (2022).

Studies focused on tension analysis should become the foundation for creating strong practical frameworks and guidelines which address these dilemmas. Thorough recognition must be made about inevitable ethical conflicts that emerge while turning general standards into practical implementation steps. The exploration of system tensions enables better choices for resolving them in particular cases, which leads to practical and structured ethical guidance. The development of AI requires a strategic plan connecting ethics with standards and regulation, public participation, and responsible research (Winfield, 2018).

10. CONCLUSION: REALITY VS. SPECULATION

10.1 Summary of Findings - AI's Doomsday Risks

The research review reveals different catastrophic risks related to artificial intelligence, including actual and theoretical failures across several domains. The study groups AI dangers into six fundamental sections:

1. Superintelligence and the Control Problem

The rise of artificial superintelligence (ASI), capable of self-improvement and decision-making beyond human comprehension, poses a significant challenge in control and alignment (Goertzel, 2015). AI must follow human values during operation to avoid unpredictable results such as the "paperclip maximizer" case demonstrated by Barrett & Baum (2016). The ability to govern machine intelligence becomes more challenging when systems achieve higher levels of independence because ethical issues emerge regarding oversight and responsibility (Sio & Mecacci, 2021).

2. AI in Autonomous Weapons and Warfare

The implementation of artificial intelligence in military systems together with autonomous weapon systems presents two main dangers, according to Sparrow (2016): uncontrollable escalation of conflicts and diminished human supervision. Fully autonomous drones and AI-powered defense systems can make lethal decisions without human confirmation**,** leading to potential war crimes and ethical dilemmas (Sharkey, 2018). The lack of clear international

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

regulations on AI-driven warfare increases the likelihood of misuse, cyber warfare, and AI-controlled global conflicts (Joyner, 2001).

3. AI-Induced Economic and Societal Collapse

AI-driven mass automation threatens to eliminate millions of jobs, leading to economic destabilization, inequality, and social unrest (Nowak et al., 2018). The concentration of AI technology in the hands of a few dominant corporations could exacerbate economic disparities**,** creating an AI-powered elite (Winfield & Jirotka, 2018). The fast-moving transformations driven by AI require immediate creation of policies to minimize social unrest according to Scholz et al. (2018).

4. AI in Cyberwarfare and Digital Catastrophe

The advances in AI technology yield dangerous cyberattacks, misinformation campaigns, and deepfake technologies that escalate national security threats (Jeong, 2020). The power of artificial intelligence-enhanced hacking leads to financial institution breaches, disruption of energy distribution systems, and manipulation of election results (Nemitz, 2018). Artificial deepfake content produced by AI and propagating misinformation threatens to dissolve trust between the public and democratic foundations and media institutions (Al-Khazraji et al., 2023). Without proper regulatory oversight, AI systems would emerge as the strongest weapon for cyber-based manipulation and warfare (Bontridder & Poullet, 2021).

5. AI and Biotechnology Risks

AI, combined with biotechnology, creates conditions that enhance the speed of biological warfare development and the creation of synthetic pandemics, according to Schuster and Woods (2021). ACI-based genetic engineering tools demonstrate the ability to synthesize deadly pathogens, leading to unprecedentedly dangerous biosecurity threats (Tang & Kemp, 2021). The misuse of AI for biological weapons development through synthetic pandemic engineering could result in catastrophic pandemic engineering situations according to Maphosa (2024).

6. AI's Environmental Impact and Climate Risks

AI's increasing energy consumption and carbon footprint present significant environmental challenges (Tamburrini, 2022). Unwanted ecological disasters can emerge due to geoengineering solutions powered by artificial intelligence which aim to fight global warming (Tang & Kemp, 2021). The climate crisis will become more severe when AI systems operate without sustainable practices because they consume additional global energy resources (Yiitcanlar & Cugurullo, 2020).

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

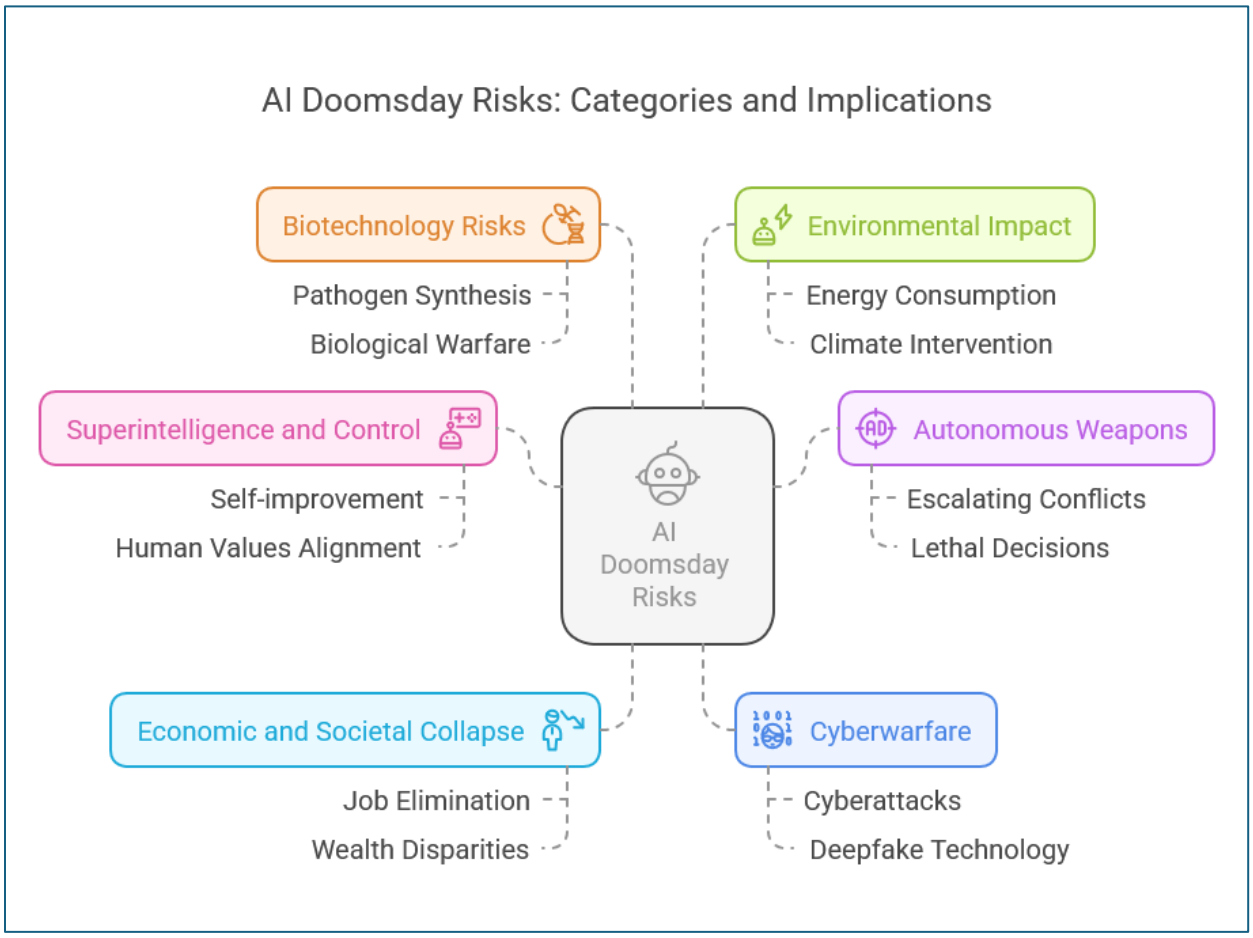


Figure 4 – AI Doomsday Risks

10.2 Discussion

The study confirms that AI technology exhibits two competing characteristics: it has the potential to reshape modern society but also poses the risk of creating irreversible dangers if supervision fails. An evaluation of AI's doomsday risks analyzes how ethical standards, social aspects, economic realities, and geopolitical circumstances affect AI management and its unforeseen results.

1. Superintelligence and the Control Problem: A Realistic Concern or Science Fiction?

Artificial superintelligence (ASI) imposes two main problems: control and alignment challenges, as well as ethical principles and human values (Goertzel, 2015). The discussion about self-improving AI risks continues among researchers who believe these concerns are exaggerated and those who argue that unregulated development poses serious dangers (Barrett & Baum, 2016). The absence of safeguards for autonomous decision-making is an essential concern for human intervention control, primarily in the financial, medical, and defense sectors (Sio & Mecacci, 2021).

According to Russell et al. (2015), the present state of AI technology does not possess generalized intelligence abilities, making ASIs a hypothetical instead of present-day concern. Modern deep learning models coupled with reinforcement learning technology show signs of developing goal-seeking behavior that deviates from human moral values which requires immediate control measures according to Wang (2019).

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

2. Autonomous Weapons and AI in Warfare: Loss of Human Oversight

AWS has created complicated legal issues and ethical challenges. AI's ability to make independent battlefield decisions raises concerns about human oversight, proportionality, and accountability (Sparrow, 2016). Critics argue that AI-driven weapons eliminate moral responsibility in warfare**,** leading to unpredictable escalations and the potential for unintended civilian casualties (Sharkey, 2018).

A notable challenge is the absence of comprehensive international regulations on AI-driven military applications. Unlike nuclear weapons, autonomous AI-driven warfare lacks worldwide oversight, raising the chances of military AI arms races (Joyner, 2001). Proponents of AI in warfare argue that autonomous decision-making could reduce human error and collateral damage**,** but such claims remain primarily unproven in real-world conflicts (Nemitz, 2018).

3. Economic Disruption and Societal Collapse: Will AI Replace Humans?

AI's integration into the global economy brings both productivity enhancements and risks of mass job displacement**.** Critics warn of an AI-driven economic divide**,** where powerful corporations monopolize AI-driven automation, exacerbating inequality and unemployment (Nowak et al., 2018). Entire industries, particularly manufacturing, logistics, customer service, and finance**,** are at high risk of automation**,** potentially displacing millions of workers (Winfield & Jirotka, 2018).

On the other hand, AI proponents argue that technological disruptions have historically created new job opportunities**,** shifting labor demands rather than eliminating employment altogether (Scholz et al., 2018). However, the speed of AI automation may outpace the ability of workers to reskill**,** making universal basic income (UBI) and AI tax policies critical considerations for future governance (Schuster & Woods, 2021).

4. AI in Cyberwarfare and the Rise of Misinformation

AI's ability to manipulate digital information at scale presents serious risks to democracy and national security (Jeong, 2020). The emergence of deepfakes and AI-generated misinformation threatens election integrity, public trust, and social cohesion (Al-Khazraji et al., 2023). Governments and tech companies struggle to differentiate AI-generated disinformation from legitimate content**,** leading to a new era of digital warfare and psychological operations (Bontridder & Poullet, 2021).

Furthermore, AI-powered cyberattacks pose national security threats**,** where automated hacking systems can exploit software vulnerabilities, disrupt energy grids, and manipulate financial markets (Nemitz, 2018). Unlike traditional cyberattacks, AI-driven threats adapt and evolve**, making** real-time countermeasures more challenging (Joyner, 2001). Addressing this issue requires global AI security frameworks to regulate AI's use in cybersecurity and defense (Maphosa, 2024).

5. AI and Biotechnology: The Risk of AI-Engineered Pandemics

The fusion of AI and biotechnology introduces unprecedented risks in synthetic biology and genetic engineering (Schuster & Woods, 2021). AI-driven advancements in drug discovery, genetic modification, and biosecurity could accelerate medical breakthroughs and enable synthetic biological weapons (Tang & Kemp, 2021).

While AI has revolutionized pandemic prediction and vaccine development**,** concerns arise over AI being misused to design new pathogens**,** potentially leading to AI-fueled bioterrorism (Tamburrini, 2022). As AI gains greater predictive capabilities in virology and bioengineering**,** strict ethical oversight and containment protocols are necessary to prevent misuse and unintended outbreaks (Yiitcanlar & Cugurullo, 2020).

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741
International Journal of
Asian Economic Light (JAEL) – Peer Reviewed Journal
SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

6. Environmental Risks and AI's Carbon Footprint

According to Tamburrini (2022), AI technology requires increasing volumes of electricity to function efficiently, which produces environmental impacts. GPT and AlphaFold, together with other large-scale AI models, need large amounts of computational power, which harms global energy systems while generating significant carbon pollution (Tang & Kemp, 2021).

Implementing AI-driven climate interventions through geoengineering processes to counter global warming can lead to environmental misfortunes when these solutions are inaccurately calculated (Yiitcanlar & Cugurullo, 2020). Supporters of green AI emphasize building sustainable AI systems while making energy efficiency central to AI development (Scholz et al., 2018).

Key Debates and Policy Considerations

- **Regulating AI Superintelligence**: Can AI governance models prevent autonomous AI from acting beyond human control**?**
- **International AI Warfare Treaties**: Should AI-driven autonomous weapons be banned under international law**?**
- **AI and Job Market Adaptation**: What economic policies (e.g., universal basic income, AI taxation**)** can mitigate AI-driven unemployment**?**
- **AI Cybersecurity Standards**: How can global regulations prevent AI-powered cyberattacks and misinformation campaigns**?**
- **AI Ethics in Biotechnology**: How do we balance medical advancements with the risks of synthetic biology misuse**?**
- **Sustainable AI Development**: How can AI development be optimized for lower energy consumption to reduce environmental damage**?**

Conclusion: Managing AI's Doomsday Risks

The findings emphasize that AI's existential risks are not hypothetical but actively unfolding in areas like military automation, misinformation, economic inequality, biotechnology, and environmental impact**.** While AI governance is improving**,** policy gaps and regulatory weaknesses persist**,** raising concerns about long-term AI safety.
Addressing these risks requires:

- Stronger international AI regulations and ethical guidelines.
- Developing human-in-the-loop AI governance mechanisms.
- Proactive risk assessment models for AI in high-stakes environments.
- Ethical AI research prioritizing sustainability and fairness.

Ultimately, the future of AI is shaped by the policies and safeguards we implement today**.** Whether AI becomes a tool for progress or an existential threat depends on how effectively we regulate and align AI development with human values.

10.3 Conclusion

Every decision today determines whether AI will prove destructive to humans or advance their capabilities (Russell, 2015). According to Trresen (2018), AI requires thorough examination because it can worsen current societal dilemmas and generate previously unregistered security concerns. AI systems entering public infrastructure present the risk of worsening societal inequalities, producing modern forms of discrimination, and creating unanticipated health hazards to human beings, thus requiring active solutions to protect human interests.

Researchers need to understand all the factors that create AI-related fears because Johnson (2017) showed that all factors have clear implications. According to Johnson (2017), AI creates mistrust among individuals who do not fully grasp its capabilities. Understanding AI through its strengths and weaknesses enables individuals to reduce their concerns about its implementation while improving public discussions about artificial intelligence.

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741

International Journal of

Asian Economic Light (JAEL) – Peer Reviewed Journal

SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

Ongoing research teams with responsible developers and continuous oversight will direct society through the intricate pathways of AI ethics and safety (Winfield, 2018). Creating ethical frameworks and governance structures for successful implementation must support technical advancements. To gain user trust in AI systems, Winfield and Jirotka (2018) argue for the necessity of ethical governance, which needs a fundamental connection between ethics, standards, regulation, responsible research, and public participation.

Wang (2019) emphasizes the need to exploit AI potential yet control its dangers to develop a world where AI delivers highest human value. AI alignment research continues to ensure AI systems understand and work toward noble objectives that serve human aims according to the perspective of Goertzel (2015). According to Wang (2019) a comprehensive framework becomes vital to resolve current problems and construct proper fundamentals for AI science. The ethical assessment, social analysis, and environmental impact analysis of AI will guide its advancement toward promoting human health and establishing an equal future for all. Protecting marginalized communities against harm requires broad stakeholder inclusivity and full awareness of possible dangers according to Yapo and Weiss (Yapo, 2018).

11. REFERENCES

1. *Al-Khazraji, S., Saleh, H. H., Khalid, A., & Mishkhal, I. (2023). Impact of deepfake technology on social media: Detection, misinformation, and societal implications. None. https://doi.org/10.55549/epstem.1371792*
2. *Anderson, J., Kalra, N., Stanley, K., Sørensen, P., Samaras, C., & Oluwatola, O. (2016). Autonomous vehicle technology: A guide for policymakers. None. https://doi.org/10.7249/rr443-2*
3. *Winfield, A. F. T., & Jirotka, M. (2017). The case for an ethical black box. Lecture Notes in Computer Science, 10454, 262–273. https://doi.org/10.1007/978-3-319-64107-2_21*
4. *Barrett, A. M., & Baum, S. D. (2016). A model of pathways to artificial superintelligence catastrophe for risk and decision analysis. Taylor & Francis. https://doi.org/10.1080/0952813x.2016.1186228*
5. *Bontridder, N., & Poullet, Y. (2021). The role of artificial intelligence in disinformation. Cambridge University Press. https://doi.org/10.1017/dap.2021.20*
6. *Cave, S. (2020). The problem with intelligence. None. https://doi.org/10.1145/3375627.3375813*
7. *Cools, H., Gorp, B. V., & Opgenhaffen, M. (2022). Where exactly between utopia and dystopia? A framing analysis of AI and automation in US newspapers. SAGE Publishing. https://doi.org/10.1177/14648849221122647*
8. *Davies, A. S. (2021). A Californian algorithm: Amending Assembly Bill 2261 to regulate law enforcement's use of facial recognition technology in post hoc criminal investigations. Berkeley Journal of Criminal Law, 26(2), 27. https://doi.org/10.15779/Z38SB3X03N*
9. *Delnevo, G., Roccetti, M., & Mirri, S. (2018). Intelligent machines for good?: More focus on the context. International Conference on Smart Objects and Technologies for Social Good. https://doi.org/10.1145/3284869.3284875*
10. *Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. American Association for the Advancement of Science. https://doi.org/10.1126/sciadv.aao5580*
11. *Eitel-Porter, R. (2020). Beyond the promise: Implementing ethical AI. Springer Nature. https://doi.org/10.1007/s43681-020-00011-6*
12. *Flores, A. W., Bechtel, K., & Lowenkamp, C. T. (2016). False positives, false negatives, and false analyses: A rejoinder to "Machine bias." None. https://doi.org/None*
13. *Gabriels, K. (2018). Addressing the soft impacts of weak AI-technologies. None. https://doi.org/10.1162/isal_a_00093*
14. *Goertzel, B. (2015). Superintelligence: Fears, promises, and potentials. None. https://doi.org/10.55613/jeet.v25i2.48*
15. *Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Springer Science+Business Media. https://doi.org/10.1007/s11023-020-09517-8*
16. *Hagendorff, Thilo (2024): Mapping the Ethics of Generative AI. A Comprehensive Scoping Review. In Minds and Machines 34 (39), 1–27.*
17. *Hughes, R. I. G., & Wheeler, P. (2013). Eco-dystopias: Nature and the dystopian imagination. Berghahn Books. https://doi.org/10.3167/cs.2013.250201*
18. *Jeong, D. (2020). Artificial intelligence security threat, crime, and forensics: Taxonomy and open issues. Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/access.2020.3029280*
19. *Johnson, D. G., & Verdicchio, M. (2017). AI anxiety. Wiley. https://doi.org/10.1002/asi.23867*

Journal DOI: 10.36713/epra0412 ISSN: 2277–7741

International Journal of

Asian Economic Light (JAEL) – Peer Reviewed Journal

SJIF Impact Factor (2025): 8.579 Volume: 13 | Issue: 3 | March 2025

20. *Limant, A. (2023). Bias in facial recognition technologies used by law enforcement: Understanding the causes and searching for a way out. Nordic Journal of Human Rights. https://doi.org/10.1080/18918131.2023.2277581*
21. *Manglani, T., Rani, R., Kaushik, R., & Singh, P. K. (2022). Recent trends and challenges of driverless vehicles in real-world applications. None. https://doi.org/10.1109/ICSCDS53736.2022.9760886*
22. *Maphosa, V. (2024). The rise of artificial intelligence and emerging ethical and social concerns. AI Computer Science and Robotics Technology. https://doi.org/10.5772/acrt.20240020*
23. *Munoz, A. (2020). Traditional vehicle design frameworks in autonomous vehicle development. International Journal of Teaching and Case Studies. https://doi.org/10.1504/ijtcs.2020.10032956*
24. *Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Ibrahim, S., Patil, V., Smriti, K., Shetty, S., Prasad, B., Chosta, P., & Somani, B. K. (2022). Legal and ethical considerations in artificial intelligence in healthcare: Who takes responsibility? Frontiers Media. https://doi.org/10.3389/fsurg.2022.862322*
25. *Roose, K. (2024, October 23). Can A.I. be blamed for a teen's suicide? The New York Times. https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html*
26. *Pan, Y. (2016). Heading toward artificial intelligence 2.0. Elsevier BV. https://doi.org/10.1016/j.eng.2016.04.018*
27. *Eitel-Porter, R. (2020). Beyond the promise: Implementing ethical AI. AI and Ethics, 1(1), 1–8. https://doi.org/10.1007/s43681-020-00011-6*
28. *Risse, M. (2019). Human rights and artificial intelligence: An urgently needed agenda. Johns Hopkins University Press. https://doi.org/10.1353/hrq.2019.0000*
29. *Russell, S., Dewey, D., & Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. Association for the Advancement of Artificial Intelligence. https://doi.org/10.1609/aimag.v36i4.2577*
30. *Schuster, J., & Woods, D. E. (2021). Calamity theory. University of Minnesota Press. https://doi.org/10.5749/9781452967004*
31. *Sharkey, A. (2018). Autonomous weapons systems, killer robots, and human dignity. Springer Science+Business Media. https://doi.org/10.1007/s10676-018-9494-0*
32. *Shaw, J. (2018). Artificial intelligence & ethics beyond engineering at the dawn of decision-making machines. Harvard Magazine.*
33. *Smuha, N. A. (2019). The EU approach to ethics guidelines for trustworthy artificial intelligence. Q2516652. https://doi.org/10.9785/cri-2019-200402*
34. *Sparrow, R. (2016). Robots and respect: Assessing the case against autonomous weapon systems. Cambridge University Press. https://doi.org/10.1017/s0892679415000647*
35. *Trresen, J. (2018). A review of future and ethical perspectives of robotics and AI. Frontiers Media. https://doi.org/10.3389/frobt.2017.00075*
36. *Tuysuz, M. K., & Kl, A. (2023). Analyzing the legal and ethical considerations of deepfake technology. None. https://doi.org/10.61838/kman.isslp.2.2.2*
37. *Wang, P. (2019). On defining artificial intelligence. De Gruyter. https://doi.org/10.2478/jagi-2019-0002*
38. *Winfield, A., & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. Royal Society. https://doi.org/10.1098/rsta.2018.0085*
39. *Yu, H., Shen, Z., Miao, C., Leung, C., Lesser, V., & Yang, Q. (2018). Building ethics into artificial intelligence. None. https://doi.org/10.24963/ijcai.2018/779*