S.Subashini, 2026, 14:4
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Artificial Intelligence: Friend or Foe

S.Subashini[1],
Assistant Professor,Anna University,Chinnai/J.P.College of Engineering,Tenkasi[1]

M.Nandhinidevi[2],
Student,Anna University,Chennai/J.P.College of Engineering,Tenkasi[2]

S.Nathiya[3],
Student,Anna University,Chennai/J.P.College of Engineering,Tenkasi[3]

K.Indhuja[4],
Student,Anna University,Chennai/J.P.College of Engineering,Tenkasi[4]

K.Madhubala[5],
Student,Anna University,Chennai/J.P.College of Engineering,Tenkasi[5]

T.Arokya Samy[6],
Student,Anna University,Chennai/J.P.College of Engineering,Tenkasi[6]

Abstract- Artificial Intelligence (AI) is a rapidly developing technology that is transforming various aspects of human life. It enables machines and computer systems to perform tasks that normally require human intelligence, such as learning, reasoning, problem-solving, and decision-making. AI is increasingly being used in education, healthcare, business, transportation, communication, and entertainment.AI offers numerous benefits to society. It can improve efficiency, reduce human effort, analyze large amounts of data, and assist people in solving complex problems. In healthcare, AI can support disease detection and medical research, while in education it can provide personalized learning experiences. These applications demonstrate how AI can act as a valuable friend to humanityAt the same time, AI creates several challenges and concerns. Automation may affect employment opportunities in certain industries, while excessive dependence on technology may reduce human creativity and critical thinking. Other concerns include data privacy, cybersecurity, algorithmic bias, misinformation, and the misuse of AI for harmful purposes.The relationship between humans and AI therefore depends on how responsibly the technology is developed and used. Ethical guidelines, appropriate regulations, transparency, and human supervision are essential to ensure that AI benefits society without causing unnecessary harm. People must also develop the skills needed to work effectively alongside intelligent technologies.In conclusion, Artificial Intelligence can be both a friend and a foe depending on its application. When used responsibly, it has the potential to improve human life, support innovation, and address major global challenges. A balanced approach that combines technological progress with human values and responsibility is essential for creating a safe and beneficial AI-powered future.

Keywords— Artificial Intelligence, technology, machine learning, automation, healthcare, education, innovation, productivity, job displacement, privacy, data security, misinformation, bias, ethics, human responsibility, regulation, benefits, risks, society, future of AI.

I. INTRODUCTION

It refers to the ability of machines and computer systems to perform tasks that usually require human intelligence, such as learning, reasoning, problem-solving, decision-making, and understanding language. AI has moved from being a concept in science fiction to becoming a part of everyday life

Today, AI is used in many different fields, including education, healthcare, banking, transportation, agriculture, business, and entertainment. AI-powered applications can help students learn, assist doctors in identifying diseases, improve business decisions, and make transportation systems more efficient. These developments show that AI has the potential to make human life easier, faster, and more productive. One of the major advantages of AI is its ability to process large .Unlike humans, AI systems can work continuously without becoming tired and can perform repetitive tasks with high efficiency. In scientific research and healthcare, AI can help researchers analyze complex data and discover useful patterns, leading to better solutions and innovations. However, Artificial Intelligence also has certain disadvantages and risks. The increasing use of automation may reduce the need for human workers in some industries, creating concerns about employment. AI can also raise issues related to privacy, cybersecurity, misinformation, and the misuse of personal data. Overdependence on AI may further affect human creativity, independent thinking, and decision-making skills. The question of whether AI is a friend or a foe therefore depends largely on how humans use it. If AI is developed with proper ethical principles, transparency, safety measures, and human supervision, it can provide enormous benefits to society. On the other hand, careless or irresponsible use of AI can create serious social, economic, and ethical problems. This study explores Artificial Intelligence as both a friend and a foe by examining its benefits, limitations, opportunities, and risks. It aims to understand how AI is influencing human life and how society can use this technology responsibly. With the right balance between technological progress and human values, AI can become a powerful tool for creating a better and more sustainable future.

II. RELATED WORK

Previous studies on Artificial Intelligence have explored its growing influence on different areas of human life. Researchers have examined how AI systems can perform tasks such as learning, prediction, decision-making, language processing, and pattern recognition. These studies suggest that AI has developed from simple automated systems into advanced technologies capable of supporting complex human activities. Several works have focused on the use of AI in education. Researchers have found that AI-based learning platforms can provide personalized learning experiences based on the needs and abilities of individual students. Intelligent tutoring systems, automated assessment, and AI-powered educational assistants can help students understand difficult topics and receive immediate feedback. However, studies also highlight the importance of maintaining teacher involvement and preventing excessive dependence on AI tools. Healthcare is another major area discussed in previous research. AI has been investigated for medical diagnosis, disease prediction, medical imaging, drug discovery, and patient monitoring. Research indicates that AI can assist healthcare professionals by identifying patterns in large amounts of medical data. At the same time, researchers emphasize concerns about patient privacy, data security, accuracy, and the need for doctors to remain involved in important medical decisions. Many studies have also examined the economic and employment effects of AI. AI and automation can increase productivity by performing repetitive and time-consuming tasks. However, research suggests that some jobs may be replaced or significantly changed as intelligent systems become more capable. At the same time, AI can create new employment opportunities requiring skills in technology, data analysis, system management, and human-AI collaboration. This has increased the importance of education and reskilling programs. Another important area of related research concerns the ethical and social challenges of AI. Researchers have discussed issues such as algorithmic bias, privacy, surveillance, misinformation, transparency, and accountability. AI systems may produce unfair results when they are trained using biased or incomplete data. Therefore, several studies emphasize the development of ethical guidelines, responsible AI practices, transparency, and human oversight to reduce potential harm. Overall, previous works show that Artificial Intelligence has both significant benefits and potential risks. Research across education, healthcare, employment, and ethics demonstrates

that AI should not simply be viewed as either a friend or a foe. Its impact depends on how it is designed, regulated, and used by society. The findings of these related works provide a foundation for further studying how AI can be used responsibly while minimizing its negative effects.

III. LITERATURE REVIEW

Artificial Intelligence (AI) has become an important area of research because of its growing influence on society. Previous studies describe AI as a technology that enables machines to perform tasks that normally require human intelligence, including learning, reasoning, prediction, problem-solving, and decision-making. The rapid development of machine learning, deep learning, and natural language processing has expanded the use of AI across many sectors. A significant body of literature focuses on the use of AI in education. Researchers have examined AI-powered tutoring systems, personalized learning platforms, automated assessment, and virtual assistants. These technologies can provide students with immediate feedback and learning materials suited to their individual needs. However, studies also point out that AI should support teachers rather than completely replace human interaction and guidance. Healthcare is another major area of AI research. Studies have investigated the use of AI for medical image analysis, disease detection, patient monitoring, drug development, and clinical decision support. Research suggests that AI can help healthcare professionals analyze large volumes of medical information efficiently. Nevertheless, concerns regarding patient privacy, data security, accuracy, and accountability remain important challenges. The literature also examines the effect of AI on employment and the economy. AI-based automation can improve productivity and reduce the time required for repetitive tasks. However, researchers have raised concerns that automation may replace certain jobs or change the skills required in the workplace. At the same time, AI is expected to create new opportunities in areas such as software development, data science, AI management, and technology-related services. Reskilling and continuous education are therefore frequently identified as important responses to these changes. Ethical issues have received increasing attention in AI research. Scholars have discussed algorithmic bias, privacy, transparency, accountability, misinformation, and the responsible use of AI. Biased training data can lead to unfair outcomes, while systems that are difficult to understand can make it challenging to determine how decisions are made. These concerns have encouraged researchers and organizations to promote principles such as fairness, transparency, safety, human oversight, and responsible innovation. Overall, the literature indicates that Artificial Intelligence has both positive and negative effects. AI can improve efficiency, support innovation, and assist humans in solving complex problems, but it can also create social, economic, and ethical challenges. Previous research therefore emphasizes the importance of responsible development and human supervision. This literature review provides a foundation for examining whether AI should be considered a friend or foe and how its benefits can be maximized while minimizing its potential risks.

1. Ethical Frameworks for Evaluating AI

Three classical ethical traditions offer complementary lenses for evaluating AI systems. A utilitarian or consequentialist lens asks whether a given deployment produces a net positive balance of benefits over harms across all affected parties — a framework well suited to cost-benefit style policy analysis but vulnerable to underweighting harms concentrated in small or marginalized populations. A deontological lens asks whether an AI system respects certain inviolable rights and duties regardless of aggregate outcomes — for instance, a right to explanation when an algorithmic decision affects someone's employment or credit access, or a duty not to deceive users about whether they are interacting with a machine. A virtue-ethics lens, by contrast, asks what habits and dispositions AI development cultivates in the organizations and individuals who build it: does a given practice encourage honesty, humility, and care, or does it reward speed and opacity at the expense of accountability? This paper does not adopt any single framework as definitive. Instead, it treats the three lenses as complementary diagnostic tools:

consequentialist analysis for weighing aggregate impact (as in Sections 3–5), rights-based analysis for identifying non-negotiable protections (as in the bias and privacy discussions in Section 4), and virtue-based analysis for evaluating institutional culture (as in the governance discussion in Section 7).

IV. THE CASE FOR AI AS A FRIEND

AI has demonstrated significant potential to improve human welfare across multiple domains. This section reviews six areas where the evidence for benefit is strongest.

1. Healthcare
Machine learning models assist in early disease detection, drug discovery, and personalized treatment planning, often matching or exceeding human diagnostic accuracy in narrow domains such as reading radiology scans or detecting diabetic retinopathy. AI-driven protein structure prediction has shortened research timelines that once took years into a matter of days, accelerating the search for new therapeutics

2. Productivity and the Economy
Automation of repetitive tasks allows human workers to focus on creative and strategic work, boosting economic output. Small businesses increasingly use AI tools for customer service, inventory management, and marketing, lowering the barrier to entry that once favored only large firms with dedicated staff.

3. Accessibility
AI-powered tools such as speech-to-text, real-time translation, and screen readers have expanded access for people with disabilities, allowing greater participation in education, employment, and civic life. Real-time captioning and translation also reduce language barriers in global collaboration.

4. Scientific Discovery
AI systems have accelerated breakthroughs in fields such as materials science, climate modeling, and genomics, helping researchers identify patterns in datasets too large and complex for manual analysis. This has proven especially valuable in climate science, where AI models improve the resolution and speed of forecasting.

5. Education
Adaptive learning platforms personalize instruction to each student's pace and needs, improving educational outcomes, particularly for students who fall behind standard classroom pacing. AI tutoring tools can provide immediate feedback that would otherwise require far more instructor time than is typically available.

6. Environmental Applications
AI supports environmental monitoring through satellite image analysis for deforestation tracking, optimization of energy grids to reduce waste, and predictive maintenance that lowers the resource footprint of industrial equipment. Utility companies increasingly use AI-based load forecasting to balance renewable energy supply, which is inherently variable, against demand, reducing reliance on fossil-fuel "peaker" plants that would otherwise be needed to cover shortfalls. Collectively, these applications illustrate that when carefully designed and deployed, AI can act as a powerful amplifier of human capability, extending what individuals and institutions can achieve.

V. THE CASE FOR AI AS A FOE

Despite its benefits, AI also introduces serious risks that warrant caution. This section reviews six major categories of harm.

1. Job Displacement
Automation threatens to displace workers in manufacturing, transportation, customer service, and increasingly, knowledge-based professions such as paralegal work, translation, and entry-level coding. The transition is rarely smooth: displaced workers often lack access to the retraining programs needed to move into newly created roles.

2. Algorithmic Bias
AI systems trained on historical data can perpetuate and amplify existing social biases in hiring, lending, and criminal justice. Facial recognition systems, for example, have repeatedly demonstrated lower

accuracy rates for women and people with darker skin tones, raising serious concerns when such systems are used in law enforcement or security contexts.

3. Privacy Erosion

Large-scale data collection needed to train AI models raises concerns about surveillance and the erosion of personal privacy. Continuous data harvesting from devices, applications, and public cameras creates detailed behavioral profiles that individuals rarely consent to in any meaningful sense.

4. Misinformation and Synthetic Media

Generative AI can produce convincing false content at scale, including deepfakes, undermining trust in media and institutions. This capability has already been used to fabricate statements attributed to public figures and to generate fraudulent evidence, complicating efforts to verify truth in public discourse.

5. Concentration of Power

The resources required to build advanced AI systems — vast computing infrastructure, proprietary data, and specialized talent — are concentrated among a small number of corporations and states, raising concerns about unequal access, market dominance, and reduced public oversight of systems that increasingly shape public life.

6. Safety and Long-Term Risk

Some researchers warn that sufficiently advanced, poorly aligned AI systems could pose long-term risks to human oversight and control, particularly as systems are given greater autonomy in high-stakes domains such as financial markets, critical infrastructure, and military operations.

These concerns are not hypothetical; many are already manifesting in current AI deployments, underscoring the urgency of proactive governance rather than reactive damage control.

7. Global Regulatory Landscape

Governments have responded to these risks with markedly different regulatory philosophies. The European Union's AI Act adopts a risk-tiered approach, imposing the strictest obligations — including mandatory conformity assessments — on "high-risk" applications such as biometric identification, employment screening, and critical infrastructure, while leaving minimal-risk applications largely unregulated.

The United States has favored a more sector-specific and market-driven approach, relying on existing agencies (such as consumer protection and employment discrimination bodies) to apply established law to AI-specific harms, supplemented by voluntary frameworks such as the NIST AI Risk Management Framework. China has pursued a state-centered model that combines rapid domestic AI investment with content-specific regulation, particularly around generative AI outputs and algorithmic recommendation systems.

Comparative overview of major AI regulatory approaches by jurisdiction

Jurisdiction	Regulatory Approach
European Union	Risk-tiered, horizontal law (AI Act) covering all sectors with binding obligations for high-risk uses
United States	Sector-specific enforcement via existing agencies plus voluntary risk-management frameworks
China	State-directed investment paired with targeted rules on generative content and recommendation algorithms
India	Emerging advisory guidelines emphasizing innovation alongside a forthcoming dedicated AI governance framework

This regulatory fragmentation creates practical challenges for AI developers operating globally, who must reconcile divergent transparency, data-handling, and risk-assessment requirements. It also creates an opening for regulatory arbitrage, where development or deployment shifts toward jurisdictions with the lightest obligations — a

dynamic that underscores the case for the international cooperation recommended in Section 8.

VI. COMPARATIVE ANALYSIS

Table 1 summarizes the friend-and-foe duality across seven key dimensions of societal impact, illustrating that the same underlying capability frequently produces both beneficial and harmful potential depending on context.

Dimension	AI as Friend	AI as Foe
Healthcare	Faster diagnosis, drug discovery, personalized treatment	Over-reliance risk, data privacy concerns in patient records
Employment	New job categories, augmented productivity	Displacement of routine and some cognitive jobs
Information	Instant access, translation, summarization	Deepfakes, misinformation at scale
Governance	Data-driven policy simulation, fraud detection	Mass surveillance, opaque algorithmic decisions
Economy	Efficiency gains, new markets and industries	Wealth concentration among a few large firms
Security	Threat detection, cyber-defense automation	AI-enabled cyberattacks, autonomous weapons

Comparative summary of AI's dual impact across key societal dimensions

This comparison underscores a recurring pattern: the technical capability itself (e.g., pattern recognition in images, or large-scale data processing) is dimension-neutral. What determines whether an application falls on the "friend" or "foe" side of the ledger is the purpose it is put to, the safeguards placed around it, and who controls its deployment.

VII. CASE STUDIES

1. Facial Recognition: Reunification vs. Surveillance

Facial recognition technology has been used by non-profit organizations to help reunite missing children with their families by matching images against databases of registered cases. The same underlying technology, deployed by state authorities without adequate oversight, has been used for mass surveillance of political dissidents and ethnic minorities, demonstrating how identical technical capability produces opposite societal outcomes depending on the governance context in which it operates.

2. Generative AI in Journalism and Misinformation

Newsrooms have adopted generative AI to accelerate transcription, translation, and first-draft summarization of complex documents, freeing journalists to focus on investigation and analysis. In parallel, the same generative capabilities have been used to mass-produce fabricated news articles and synthetic images designed to mislead the public, illustrating the dual-use nature of large language models and image generators.

3. AI in Hiring: Efficiency vs. Discrimination

Automated resume screening tools can process thousands of applications far faster than human recruiters, reducing time-to-hire and administrative cost. However, several widely reported cases have found such systems downgrading candidates based on proxies correlated with gender or ethnicity learned from biased historical hiring data, resulting in discriminatory outcomes that would be illegal if performed explicitly by a human recruiter.

4. Autonomous Systems in Transportation

Advanced driver-assistance and autonomous vehicle systems have demonstrated the potential to reduce accidents caused by human error such as fatigue or

distraction. At the same time, high-profile incidents involving autonomous vehicles have raised questions about liability, testing standards, and the readiness of regulatory frameworks to keep pace with deployment.

5. AI in Financial Services: Inclusion vs. Exclusion

AI-driven credit-scoring models have expanded access to financial services for individuals without traditional credit histories by incorporating alternative data such as utility payments, extending inclusion to previously underserved populations. Conversely, opaque scoring models have also been shown to encode proxies for protected characteristics, in some cases resulting in creditworthy applicants being denied loans for reasons that would be difficult to identify or contest without algorithmic transparency requirements.

VIII. BEYOND THE BINARY: A TOOL SHAPED BY HUMAN CHOICES

The friend-or-foe framing, while rhetorically useful, oversimplifies a more nuanced reality. AI itself is not an autonomous moral agent; it is a set of tools and techniques whose consequences depend on who builds them, how they are trained, what objectives they optimize for, and the regulatory environment in which they operate. The case studies in Section 6 make this concrete: in each pairing, it is not the technology that changes between the beneficial and harmful outcome, but the surrounding institutional context — oversight, consent, transparency, and accountability.

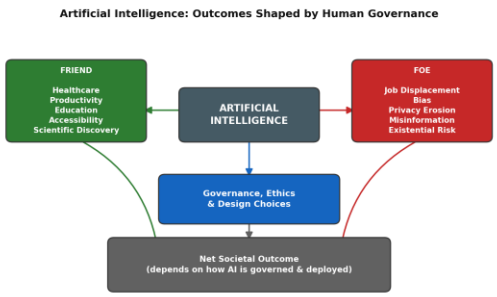

Figure 1: AI outcomes are shaped by governance, ethics, and design choices, not by the technology alone.

This suggests that the central question is not whether AI is good or bad, but how societies choose to govern its development and deployment. A governance-centered lens shifts the debate from abstract technological determinism toward concrete, actionable questions: who is accountable when an AI system causes harm, what transparency is owed to those affected by algorithmic decisions, and what institutions are best positioned to enforce these standards at the pace AI development demands?

Towards Responsible AI: Recommendations

Regulatory and Policy Measures

- Establish clear regulatory frameworks that mandate transparency, accountability, and auditability of high-stakes AI systems.
- Require impact assessments before deployment of AI in high-risk domains such as healthcare, criminal justice, and employment.
- Encourage international cooperation on AI safety research to address risks that cross national boundaries.

Technical and Organizational Measures

- Promote diverse and representative training data, along with regular bias audits, to reduce discriminatory outcomes.
- Strengthen data privacy protections and give individuals meaningful control over how their data is used to train AI systems.
- Build human-in-the-loop review for consequential automated decisions rather than fully autonomous decision pipelines.

Workforce and Education Measures

- Invest in workforce retraining programs to help workers transition into roles that complement rather than compete with automation.
- Foster public AI literacy so that citizens can critically evaluate AI-generated content and participate in governance debates.
- Integrate AI ethics education into technical curricula so future developers internalize responsible design practices early.

Future Outlook

Looking ahead, three trends are likely to shape the trajectory of the friend-or-foe debate. First, AI

capabilities will continue to generalize across tasks, increasing both the potential benefits and the stakes of poor governance. Second, regulatory frameworks — still nascent in most jurisdictions — will mature, though likely unevenly across regions, creating a patchwork of standards that global AI developers must navigate. Third, public AI literacy will play an increasingly decisive role: informed societies are better positioned to demand accountability and to distinguish legitimate innovation from harmful deployment. None of these trends guarantee a favorable outcome. They instead point to a future in which the friend-or-foe balance remains an ongoing, contested process rather than a question that is settled once and for all. A further consideration is the pace mismatch between technological deployment and institutional adaptation. Historically, major technologies such as electricity, automobiles, and the internet took decades to be matched with mature safety standards, licensing regimes, and liability frameworks. AI's deployment cycle is considerably faster, compressing the window in which society can observe harms, build evidence, and legislate in response. This suggests that anticipatory governance — regulation informed by foreseeable risk categories rather than only by harms already observed — will be more necessary for AI than it was for earlier general-purpose technologies. Finally, the role of civil society and independent research should not be understated. Investigative journalism, academic audits, and whistleblower disclosures have repeatedly been the mechanism by which algorithmic harms first came to public attention, often preceding formal regulatory response by years. Sustaining independent capacity to study AI systems — including meaningful researcher access to model behavior and training data provenance — is therefore itself a governance priority, not merely an academic convenience.

IX. CONCLUSION

Artificial Intelligence is neither an unambiguous friend nor an inevitable foe. It is a powerful and rapidly advancing technology whose ultimate impact on society will be determined by the choices made today by researchers, corporations, policymakers, and citizens. The evidence reviewed in this paper — spanning healthcare, employment, information ecosystems, and governance — shows that identical technical capabilities can produce dramatically different outcomes depending on the institutional context in which they are deployed. By pairing continued innovation with robust ethical safeguards, transparent governance, and inclusive policy-making, it is possible to steer AI development toward outcomes that broadly benefit humanity while minimizing its very real risks. The task ahead is not to resolve the friend-or-foe debate philosophically, but to build the institutions and practices that ensure AI remains accountable to human values.

REFERENCES

1. Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
2. Floridi, L. (2019). The Ethics of Artificial Intelligence. Oxford University Press.
3. Crawford, K. (2021). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.
4. World Economic Forum. (2023). The Future of Jobs Report.
5. OECD. (2023). AI Principles and Policy Observatory.
6. Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research.
7. Bostrom, N. (2014). Superintelligence: Paths, Dangers, Strategies. Oxford University Press.
8. European Commission. (2024). EU Artificial Intelligence Act: Summary and Implementation Guidance.
9. Jumper, J., et al. (2021). Highly Accurate Protein Structure Prediction with AlphaFold. Nature.
10. Zuboff, S. (2019). The Age of Surveillance Capitalism. PublicAffairs.
11. National Institute of Standards and Technology (NIST). (2023). AI Risk Management Framework (AI RMF 1.0).
12. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and Machine Learning: Limitations and Opportunities. fairmlbook.org.

13. Vinuesa, R., et al. (2020). The Role of Artificial Intelligence in Achieving the Sustainable Development Goals. Nature Communications.
14. Acemoglu, D., & Restrepo, P. (2019). Automation and New Tasks: How Technology Displaces and Reinstates Labor. Journal of Economic Perspectives.