Deep Science Publishing, 2025
https://doi.org/10.70593/978-93-7185-753-6_7

Chapter 7: Artificial Intelligence Ethical and Societal Impacts

Swarup Panda
SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

1. Introduction to AI Ethics

The tremendous societal and economic promise of AI has been emphasized repeatedly of late. Ubiquitous applications of AI permeate the fabric of every society, discuss every facet of our future and consider every sensitive socioeconomic issue [1-2]. Societal sectors such as health and medicine, transportation, communications, infrastructure, education, defence, governance, environment and others use, interact with and deliberate on the benefits and drawbacks – sometimes specific, sometimes general – of AI. These applications create many benefits, but also create certain ethical challenges, dilemmas and concerns. Industry developers aspire to produce autonomous systems that can make ethical decisions. Much work has considered how humans might produce such ethical systems.

Despite hopes to the contrary, the recent advances in AI have succeeded so well in mimicking human capabilities that it has become difficult to distinguish AI applications from human-based applications. These advances have led to not only many proposals for new AI capability, but also exams of UN resolutions for the development and use of new AI technologies under recent controversial geopolitical activity. New UN assessment activities provide guidance aimed at protecting the safety, civil liberties and privacy of citizens who interact with national military use of AI technologies. A related group of countries has recently proposed a UN resolution to ban any use of AI technologies that can replace humans in the authority to launch a nuclear strike. Ultimately, society must face new dilemmas about how much discretion it is willing to give to autonomous AI-enabled tools in such sensitive and critical areas.

2. The Importance of Regulation in AI

The rapid growth of the AI field comes with a strong motivation: Regulatory bodies, governments, and society at large are also taking a special interest in the application of these new technologies. Everyone agrees that regulations for algorithms and data are necessary, but the question is how and when. Ethical principles about AI are helpful in prioritizing data policies. For example, fairness in AI relates directly to diversity in data, while the transparency principle is associated with the use of legal and use-only-consented-for data.

Although many concerns about the risks of AI are related to factors that are external to the ML system, such as a usage context that makes the system prone to risky decisions, often the root cause can be traced to the data the model was trained on. Counters to these risks can therefore also be handled—at least partially—through data-focused measures. Ethical principles can provide guidance as to what those measures should be, though these principles do not come with a strict scorecard or clear set of evaluation metrics. Instead, using ethical principles is about evaluating broadly whether a decision aligns with a particular principle; it is not about calculating a decision's fine-grained fairness or transparency value on a scale from 0 to 1. Nonetheless, within the context of ML implementation, it is possible to move towards a type of scoring system—one based on fairness or transparency in a narrowly defined research question.

2.1. Current Regulatory Frameworks

Recent advances in artificial intelligence (AI), robotics, and automated means of implementing computational intelligence have raised. The development of new technologies, innovations, and the resulting implications challenge our established legal ethical and societal norms [3-5]. A number of concerns have been addressed by the current set of policy and regulatory structures, such as data privacy, intellectual property, and the mitigating societal impacts witnessed. These regulatory bodies have been established to protect the interests of society, safeguard civil liberties and individual freedoms, as well as to address issues like national security and crime prevention.

Though these structures have served as an initial guide, they are inherently reactive in nature, having been designed for structures other than AI and governance. New technologies can learn or evolve—much faster and with less bias compared to humans—making certain regulations redundant or inadequate. Moreover, the emergence of new offensive and novel applications introduces types of harm not addressed by the current policies or jurisprudence. Data privacy presents an ever-present challenge, especially in the context of large-scale data breaches. Robust security measures are essential to prevent such breaches and address concerns regarding the algorithmic transparency of AI and associated fairness and benefits distribution.

2.2. Proposed Regulations for AI

Among the many proposed regulations, the ChatGPT Regulation Act presented in the NSW Parliament calls for, among other things, that AI companies disclose the source of their training data for AI chatbots, that AI-generated content be clearly marked as such, that users of AI chatbots be warned about the risks of inaccurate information and misinformation, that risks relating to the protection of users' privacy be disclosed by the developers of AI chatbots, and that there be clear procedures by which any content generated by AI chatbots which is intended to defame or harass any person can be promptly removed.

Another prominent proposal was fashioned by the Co-Chairs of the Foundation Model Task Force, who offer materials to support bipartisan Federal legislation applicable to foundation-model providers that focuses on the risks of Foundation Models, such as ChatGPT, DALL-E, and Stable Diffusion [6-8]. The proposal aims to ensure that foundation-model providers are responsible for taking reasonable measures to mitigate the risk of foundation models and inform the public about residual risk, and that risks warranting special caution receive special regulatory attention. Their proposal focuses on foundation-model providers because foundations-models have the largest and most systemic risks, risk mitigation demands the greatest expertise and resources, and their task force engages more directly with foundation-model firms (including OpenAI, Meta, and Anthropic). Indeed, Foundation-Model is already the subject of a recent Directive proposed for regulation at the EU level.

3. Fairness in AI Algorithms

Several popularly-deployed AI applications deserve consideration from the perspective of group harms or at least strong concerns about subgroup-centric balance/equity. One category of AI algorithms for which concerns about fairness have long been expressed is predictive algorithms. Finally, even nonpredictive algorithms can raise fairness concerns, as in the example of AI systems for game play. Specifically, it needs to be understood how these algorithms are making recommendations or using the information to make decisions. Consider a seven-category taxonomy of concerns about the fairness of predictive algorithms:

Predictive algorithms are controversial in all the three abovementioned societal domains. Consider, first, the use of predictive algorithms in banking or law/enforcement. As an example, predictive algorithms are frequently used in banking—in particular, to determine an individual's credit worthiness. The four categories in which predictive algorithms are used in this space are: credit, employment, education and housing. Predictive algorithms for credit are designed to classify individuals by their ability to pay back a loan; for employment, they assess a candidate's job fit or firing risk; for education, they classify students in terms of their needs or college success; for housing, they assess risk of tenancy or predict rental value. They are deployed for factual purpose such as: screening prospective employees, facilitating charitable giving, enhancing prostitution, providing legal advice about bail decisions, collating forensic evidence, detecting violence in online art and graffiti, and screening art work for toxicity.

3.1. Understanding Bias in AI

Recognizing and managing bias is paramount as artificial intelligence systems become more widespread. Defining bias requires specifying a particular aspect of the system and why it is being considered. The relationship is tightly bound: an AI system judged on a given aspect is biased if it systematically and unfairly discriminates against specific individuals or populaces. An imbalanced corpus of training data, typically favoring privileged groups, engenders models that accentuate bias and discrimination towards under-served sectors of society. Consequently, a reliable set of metrics for identifying bias in AI systems is necessary.

The systems approach employed by Bowen is especially valuable for AI harm and bias analysis. Engaging stakeholders outside of computing provides essential perspectives that shape the direction of analysis and elucidate the broader societal effects of bias. Abebe advocate for bias and fairness analyses that transcend the technology, extending into its interplay with the world and society. Key questions include the requisite scope and scale of bias detection and mitigation, and the causes underpinning discriminatory behaviors in AI.

3.2. Techniques for Ensuring Fairness

Algorithms can reflect biased behaviors learned by models from biased historical data or stigmatization when they generate labels for training datasets [7,9-10]. For example, an algorithm that learns from past police records to predict crime hotspots might direct increased patrols to minority neighborhoods, perpetuating higher crime and arrest rates there. Other sources of unfairness include measurement bias—errors in data for few classes or due to variations in facial expressions and poses—and systemic bias arising from structural inequities in society. These sources of algorithmic bias can sometimes be corrected through standard data-preprocessing techniques, such as reweighting and resampling.

Bias mitigation approaches during training involve adding fairness constraints to the objective function or learning unbiased data representations. Postprocessing methods include modifying the outcomes to reduce discrimination. Although such methods can reduce bias, the use of multi-criteria optimization sometimes results in weaker performances in other measures. Model explainability techniques can complement fairness by highlighting discriminatory criteria or decisions. One tactic is to focus on interpretable models, which are usually simple and generally provide an interface that helps humans reasonably understand their decision-making process. Another tactic is to develop methods for explaining complex models, including feature importance-based methods, example-based methods, and model internal structure-based methods.

3.3. Case Studies of Fairness Failures

Bias mitigation efforts are often oriented toward surrogate goals and partial-completion metrics. Few interventions have proven effective at decisively producing fairer allocations in practice. Prolonged application of such partial-completion methods can induce measure-specific overfitting, resulting in relatively small score differences, but without any meaningful advance in fairness or social equity. The presence of confounding, pre-coded biases, the relative ease of inducing artificial "fairness" improvements in complex, real-world datasets, and negative correlations between popular representational and allocation fairness metrics all render intrusive statistical fairness adjustments unfit as standalone tools for enhancing social good within consequential operations.

List-based recommendation constitutes one highly consequential real-world domain that lacks robust methods for automated fairness improvement [1,11-14]. When a series of items for multiple controversies is recommended, biasing algorithms can induce statistically significant alterations in the rankings of over 80% of the items. This procedure can yield a dedicated fairness metric, mitigating representational harms in controversial topics, so that resource exposure is more equitably allocated—allocative fairness.

Auto-completion services present an additional case in point. The property of ordered list fairness requires that each element of the output list be generated in a fair manner concerning the demographic characteristics of the user. Fairness adjustments, pursued exclusively through pre-processing, in-processing, or post-processing approaches to machine-learning task models, result in unsuitable combinations with the inherent list-generating nature of auto-completion. To achieve ordered list fairness, a novel framework is required, encompassing a user attribute inference module, a prefix-to-suffix-generation model, and a bias mitigation algorithm that reorganizes suffix candidates for each prefix into an output list that is both credible and fair.

4. Data Privacy Concerns

Massive data collection can reveal almost anything about a person, and technologies such as facial recognition give governments and companies the ability to track people across times and places. Criteria-Based Reporting or Classified Analysis Reports with geolocation can be used to protect privacy. Classification by CIA standards – Top Secret, Secret, or Confidential – helps ensure that only the right people have access to sensitive information about places where people live, such as an address or geocode.

4.1. Data Collection Practices

Before the era of Artificial Intelligence (AI), socio-structural biases already tainted the; hence largely subjective and typically inaccurate, reasoning of humans. With the recent advances in computational power, algorithmic methods became able to take over much cognitive function and decision-making. This automatic reasoning is however fed with several social biases perpetuating discrimination and unfairness by proxy of past data.

Users concern about privacy is rising worldwide in reaction to the big data collection practices of modern internet services. A study suggests that the use of cookies, advertisements and access to phones contacts are seen as reasons to avoid digital platforms. The European Union has enacted the General Data Protection Regulation (GDPR) to enforce data protection and privacy. Its call for fairness includes the right to an explanation about an algorithmic decision. The desire to be treated equally in algorithmic decisions even led to the formulation of group anti-discrimination principles.

4.2. User Consent and Transparency

User consent and transparency. In an ideal system, a user being subjected to AI systems can always decide beforehand whether they want to be, where and how their data may be used, and for what specific purpose. Consent must be informed, freely given and easily revocable at every step in the process of creation, implementation and interaction with AI algorithms. In practice, users often accept opaque terms-of-service to quickly proceed towards the intended use of the platform, unaware of how their data will be collected and used. Inevitably, user data collected through AI applications is often treated as a common resource, and used for the training, validation and auditing of other systems, with or without explicit consent [13,15-17].

Transparency refers to the visibility and foreknowledge of all processes, parameters and procedures employed in the development and implementation of an AI system, such that the disclosure of information enables scrutiny by any stakeholder. Transparency is desirable to ensure that the system is operating in accordance with its stated goals, and is closely related to the concepts of interpretability and explainability. Currently, with deep neural networks demanding enormous complexity in their creation and implementation, the decisions made by such systems at the point of inference are often not explainable, even by their creators.

4.3. Impact of Data Breaches

Generally speaking, data breaches can impact diverse groups of users depending on the target, efficacy, and reach of the attack [18-20]. An attack on a large social media company, such as Facebook or Twitter, could impact everyday users in the form of leaked sensitive information, privacy violations, unwanted and unsolicited communication from bad actors, or even a social media identity theft.

Social media identity theft stems from the fact that when we consider social media, the whole picture is not addressed or tackled, namely that the platform can be abused for an „evil“ purpose through the creation of fake accounts, profiles and identities. This specific birth of an evil twin represents a first step in order to mask or anonymize the real content provider and onset the victims. Usually targeted attacks are initiated from such accounts in which the main objective aims at fraud and deception.

5. Societal Impacts of AI

Far from being purely hypothetical, all of the above raises important societal issues. Popular and press treatment have been started, but, faire plus ample développer, one might consider the entire landscape of impacts, from those on economy and security to those on individual freedoms, social fabric, sand human nature itself. Previously, it was suggested that many jobs were largely immune to automation, and that AI would probably be a job-creation force—hence a spur to the economy rather than a force for stagnation or decline. However, a burgeoning field of computational economics suggests that the verdict is much less clear: further automation could indeed undermine aggregate demand . In a similar vein, the market's promise to provide accelerated growth during the final stages of industrialization is something of an untested hypothesis that may well yet be disproved.

Addressing questions of security and justice, it is important to recognize that AI is vulnerable to all the classic perversions of computer applications. Thus, intrusion on privacy may be more effective than ever [19,21-22]. Digital property is vulnerable not only to glitching and sabotage, but also to hacking and theft. Getting the AI justice service right requires that we, as a society, be willing to multistage the AI justice systems ahead of time and to bear the attendant risk of flawed rulings, thereby paying initial imperfection for eventual equity. Beyond the violations of privacy and justice that can be aided and orchestrated by AI systems themselves, a more subtle concern is that the ever-increasing reliance placed on these technologies as institutions of social governance, justice, and reprisal, together with the concomitant transfer of responsibilities away from humans, may give rise to criminal and terrorist actions whose nature is fundamentally different from, and potentially more damaging than, previous generations.

5.1. AI in Employment

How does the use of AI impact employment? This question is receiving significant attention, given the public fear of robots replacing humans in the job market. Unlike the One Hundred Year Study on Artificial Intelligence (AI100), the United Nations (UN) chose to focus on AI and employment in the Workplace Learning for the Future discussion featuring Marjorie Scardino, Leona Achtenberg, Larry Cornett, and James Manyika in the Economic and Social Council (ECOSOC) Chamber at UN Headquarters in New York on 5 May 2020.

Making us all better at using AI in employment necessitates considering two dimensions of practitioner education: workplace learning and building a culture for learning. The five World Economic Forum centers are educating millions of people on AI—many professional women among them. IBM Institute for Business Value research suggests that organizations with a culture for learning (where AI finds a ready application) are three times better positioned for the future. Excelling at public engagement is essential to raising awareness about AI's potential for good and reducing fear—demonstrated in China's Health Code application, which helped restart the country's economy by reducing the risk of COVID-19 infection. Leading through times of disruption means having the same taught, caught, and sought disconnect as is seen in others. The Society, Ethics, and Policy Impact Gateway raises awareness about a human-centered approach to AI and Worker Safety.

5.2. AI and Social Inequality

Social inequalities have often accompanied major technological development and, as a consequence, public debate on new AI techniques has recognised the potential for unfair bias. Legal initiatives in the United States and Europe focus on the application of well-known anti-discrimination laws to AI. Despite the usefulness of such efforts, ultimately much larger scale structural inclusion policies will be required to avoid the exacerbation of social inequalities by the advance of AI technologies [11,23-25].

Bias in AI applications is one of the issues at the intersection of AI and the legal domain. The standard legal framework to understand the notion of bias usually arises from specific offenses of discrimination related to race, ethnicity or sex. Apparently, it would suffice to test any statistical model by verifying that there are no significant differences in the confidence intervals of a given prediction for different classes of protected attributes; for example, checking that a classifier does not generate a list of job candidates for a company, biased by sex.

5.3. Public Perception of AI

The public perception of AI—that is, how people believe AI functions, and what concerns and hopes they have about it—is very important for several reasons. When new technologies become prominent within a society, a great deal of attention focuses on them, often accompanied by a certain amount of apprehension and anxiety. People worry that humans will lose their jobs and that AI will become a threat to human life. For example, when the first steamboat was invented, some Christian groups worried that it would reverse the flow of the rivers and bog the engine down. When early computers were built, they were labeled as flame-spitting, carbon consuming, solder-solder-screaming monsters. These fears and anxieties are not new nor are the public responses to them, which generally consist of laughter, humor, cartoons, jokes, and comedic reels.

The public not only informs technology policies in a democratic society but also provides valuable insights on the ethical considerations involved with an innovation [26-28]. Some public fears and aspirations might be ‘‘overblown,’’ but they usually open the door for a careful examination of the likely consequences of an innovation. To ensure that AI never solves its proposed problems at the expense of the larger issues that faced humanity, the public must participate in the discussion.

6. Ethical Frameworks for AI Development

Objectivism is presented as a forecasting method designed for optimizing investments in disruptive technologies, specifically AI. Unlike detailed ethical or philosophical frameworks, it addresses a pragmatic embodiment of ethical principles using AI technology. Objectivism arises under the assumption of the inevitability of AI development—both inescapable and unstoppable—within the current socioeconomic paradigm (encompassing the evolving dynamics of society, economy, and its participants). A direct consequence of these assumptions is that similar superintelligent AI systems will emerge globally and become competitive, culminating in what proponent Tom Varoglu refers to as an "AI intelligence catastrophe," a concept deeply wired into forecast structures within the Objectivism framework.

Mathematical models disclose a decisive moment in the evolution of superintelligent AI that may necessitate extraordinary measures and actions. The framework's analysis delineates conditions that could compromise security guarantees of the upgrading process, plausibly leading to accelerated technology proliferation and deployment. Moreover, various solution approaches are concisely outlined. These evaluations highlight the pertinence of Objectivism-oriented thinking when applied to steering AI development along safe and socially beneficial trajectories.

6.1. Utilitarian Approaches

Two main strands of AI ethics have emerged, resulting from the different perspectives taken. The first strand—functionalist, utilitarian, rational choice, and consequentialist—focuses on how to allocate resources, maximize output, and minimize risks and externalities. It draws on AI methodologies and quantifies the impact of AI on people and groups with respect to happiness, economic well-being, wealth, and other measurable outputs. In business terms, such considerations revolve around the core philosophy of creating human, economic, and social value, based on the Aristotelian concepts of eudemonia and aretē (excellence). / Given a set of values and a model of the world, decision theory prescribes which actions will maximize the adequacy or expected performance of outputs. It answers the question, What should the AI system do, given that it must be consistent and self-consistent? The choice of inequality axioms in the design of the social welfare function affects how properties such as equality, equity, and fairness are encoded and then measured in terms of the effects on happiness or well-being. From a consequentialist perspective, Kantian deontology (immanence of ethics) can be overridden by utilitarianism, based on the proposition that we should always act so as to maximize happiness overall, but decisions must have consequences; ethics without teleology is just morality. This view has therefore been dubbed the "angry Kantian response to utilitarianism," viewed as an amoral calculus of the trade-off between harms and benefits, with little concern for the meaning of being human.

6.2. Deontological Perspectives

A consequentialist approach in ethics ranks in order of importance the possible positive and negative consequences for all humanity, enforcing a moral action when the aggregate is positive [29-32]. Whereby the deontological position is the reciprocal in that it requires all behaviour to be held to the same test hence requiring all of the particular cases and responses within the global AI discussion to be addressed adequately and properly. This makes sense since there can be no set of consequences encompassing everyone equally and is the root of many other moral theories such as Marist thought (United Nations Publications, 1993; Feinberg, 2017).

The deontological system is based on a few impenetrable principles. The UN Declarations (United Nations, 1966, 1968, 1976, 1986, 1993) on human rights and civil, political, economic, social, and cultural rights are likely candidates for such principles and are difficult to contradict. Everything about AI then revolves around these principles and whether or not it accords with them at every point of instantiation. This was suggested by Asimov in the Three Laws of Robotics (Asimov, 1950), and more fully stated in Anderson and Anderson (2011). These laws are: The First Law, A robot may not injure a human being or, through inaction, allow a human to come to harm; The Second Law, A robot must obey the orders given it by human beings except where such orders would conflict with the First Law; and The Third Law, A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

6.3. Virtue Ethics in AI

The application of ethical principles to Artificial Intelligence (AI) has been approached from deontological, utilitarian, and virtue ethics perspectives. Virtue ethics emphasizes the importance of developing traits such as courage, compassion, and wisdom. AI would act ethically by developing a virtuous character that allows it to, for example, perform the kinds of actions an ideal role model might perform in the circumstances. From this standpoint, the virtue of professional responsibility has been suggested for developers, given they have an obligation to produce AI that benefits society.

A perhaps more direct application involves formulating virtue ethics for AI. Adaptation of the principle of utility within that domain was suggested in the 1960s through an analysis of the calculus of pleasure and pain performed by Jeremy Bentham. He advocated an algorithmically based method for decision making, which can be seen as early precursor of the ethical dimension of AI. Algorithmic processes then became widely adopted for business, advertising, marketing, and most other areas of the economy. With the advent of big data, smarter algorithms can be formed, but they still suffer from the lack of ethics.

7. International Perspectives on AI Ethics

A survey and analysis of international regulations on AI, focusing on the Union of European Football Associations (FIFA) in comparison with the European Union. Internet users expressed concerns such as the excessive use of personal data collected by companies, lack of transparency about data usage, absence of coherence and clarity in use policies, and suspicion that such data are used for state control in certain countries. As long as progress remains unregulated, artificial intelligence may generate discrimination in decisions on job hires, home mortgages, judicial sentencing, health insurance, and creditworthiness assessments.

The ethical and societal impacts of artificial intelligence are widely debated. Ensuring the ethical use of AI in daily life requires agreeing upon a set of minimum ethics standards. NASA's 2019 principles for the ethical use of AI are suggested as a basis for building a common code of ethics, which should be embedded in legislative frameworks. International corporate approaches to AI ethics are also examined, complemented by analyses inspired by theories of Utilitarianism and Aristotle's Virtue Ethics.

7.1. Comparative Analysis of Global Regulations

Recent years have seen a growing number of proposals for rules and regulations affecting AI systems [31,33-35]. Much of the rationale underlying these proposed guidelines highlights issues of fairness for users, as well as data privacy concerns. Governments in some jurisdictions have taken initial steps toward regulating AI, through legislation or the introduction of specific strategies, although definitive, cross-sector, and comprehensive frameworks are not yet in place.

While the different regulations share key goals and areas of focus, agencies in other countries are likely to implement AI rules within the framework of their own national law and with respect to local conditions. Industry practitioners and companies can examine the overlap between different regulatory matrices and, in particular, compare best practices against the location of their own users and clients. Beyond regulation, there are many other aspects of the ethical and social impacts of AI that warrant specific analysis, among them transparency, trust, responsibility, safety, and copyright, as well as carefully considered theoretical approaches to ethical AI.

7.2. Cultural Influences on AI Ethics

Cultural values exert a profound influence on the ethical development of artificial intelligence, the framing and perception of relevant regulations, and the appreciation of its societal impacts. The millennia-long tradition of Confucian philosophy has helped shape Chinese cultural values. Confucian culture favors community-oriented and incremental innovation: a promotion and protection of the community has precedence over freedom of expression and privacy protection.

Some studies postulate the adoption and acceptance of a paternalistic AI model in China, where AI technologies are both promoted and be used for social governance. Other research identify protective practices in place to prevent the negative impact of artificial intelligence during times of crisis. Similarly, in the United States, where the economy is heavily influenced by the pioneer spirit, public perceptions have hinged on notions of combat and conquest. U.S. public opinion characterized robots as powerful, brave, male, direct, and cruel.

8. Future Directions in AI Ethics

Eight ways forward to help avoid Harms in AI deployments The intent here is to suggest approaches for technologists, business and public policy decision-makers, and researchers at the application stage to advance meaningful progress toward AI deployments and outcomes that are beneficial rather than harmful to individuals and society. Many of the suggestions emerge from the societal Harms framework, the consequences of ignoring the warnings and cautions therein, and the Principles for Ethical AI Selection, Adaptation and Deployment [36-38]. The suggestions look beyond a single principle, toward the interrelationships among the sets of principles that emanate naturally from the societal Harms framework, and on through the questions aimed at meaningful stakeholder consultation. The last two suggestions are especially intended for AI technologists: achieve a better understanding of the application in which the AI method or tool is to be deployed; and maintain a broad sensitivity to the multiple and varying potential Harms that can arise in the deployment of any tool or methodology. The six intermediate suggestions follow from an understanding of the relationships among three sets of ethical AI principles, each of which in turn arise from the societal Harms framework.

8.1. Emerging Technologies and Ethical Considerations

Emerging technologies affect societies, environments, and the ethical fabric of global and local communities [1,39-41]. While emerging technologies can foster innovation and drive societal progress, they also raise significant ethical and social impact issues. Moreover, many of these technologies—which include nanotechnology, biotechnologies, robotics, artificial intelligence, and cognitive sciences—are deeply interconnected. The ethical nature of a specific emerging technology is largely derived from the way it is integrated within society as a whole.

Advancements within and across these areas promise benefits such as the prevention of diseases and the advancement of the human condition [42-44]. However, they also raise concerns, including the potential suppression of innovation-based competition through patenting strategies, increases in economic inequality, contamination of food and human bodies, encroachment upon individual privacy, creation of agents capable of performing tasks that threaten the economic livelihood of workers, and the emergence of new forms of control over individuals' cognitive existence.

8.2. The Role of Stakeholders in Shaping AI Ethics

Communities, governments, regulatory bodies, investors, customers, academic institutions, and policymakers all play vital roles in shaping AI ethics. The development of applied ethical frameworks demands a comprehensive methodology grounded in the contributions of multiple stakeholders. Broad discourse on applied ethical principles must be considered throughout AI research and implementation life cycles. Such discussions have practical consequences for AI development, for example in defining governmental regulations addressing AI-related social risks. Establishing AI governance, management, monitoring, and evaluation frameworks also relies on such collaborative input.

Research into AI ethics policy assessments includes systematic taxonomies, evaluations, and comparisons of major AI principles. Surveying awareness of badly behaving AI in areas such as chatbots and recommendation systems is also necessary for constructing policy approaches [45-46]. Examples of proposed frameworks include establishing responsibilities for the consequences of AI behavior, conducting regulatory impact assessments in the context of GDPR, organizing AI principles based on established guidelines, and verifying policy assignment and evaluation for Artificial General Intelligence [18,47-49]. Furthermore, investigations into ethical and social risks associated with vulnerable groups contribute to the formulation of enforceable safeguards.

9. Conclusion

The control of intelligent machines by government or private organizations raises the question whether rules and regulations are needed to maintain control. Otherwise, ignoring this could cost many jobs, and a strong inequality gap may develop in society. Many of today's job functions, such as the delivery profession, are within close range of being replaced by autonomous vehicles and drones. To help organizations shape policy around AI, six basic elements must be included: transparency, security and quality, accountability, privacy, protection and well-being of human beings, and finally, psychological and moral well-being.

The disclosure of AI ethics and information may be protected by current intellectual property laws and regulations. AI-powered clips widespread across the Internet make attribution difficult, and the copyright of AI may be questioned in the future. This prompts questions such as “Should AI-generated works be protected by copyright? If so, who will be the copyright holder?” To ensure reliability, quality, and safety, AI should be tested with an accountability lens beyond technical elements, including data quality, security, and embedded bias. Key questions include “Who is responsible for errors, distortions, or biased decisions that are caused by AI?” and “What should the organization do if AI undermines its reputation?” The Protection of Personal Information Act (POPIA) promotes responsible information processing practices, but the right to be forgotten is difficult to apply. With the protection and well-being of humans in mind, governments and private organizations must design policies to prevent harmful effects on society. Employment may become increasingly rare, and the risk of psychological problems for the elderly might rise due to the lack of human contact. Although psychological and moral well-being are often neglected in policymaking, a culture of care should be encouraged for both AI creators and users.

References:

[1] Whittlestone J, Nyrup R, Alexandrova A, Dihal K, Cave S. Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. London: Nuffield Foundation. 2019:1-59.

[2] Vesnic-Alujevic L, Nascimento S, Polvora A. Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks. Telecommunications Policy. 2020 Jul 1;44(6):101961.

[3] Qian Y, Siau KL, Nah FF. Societal impacts of artificial intelligence: Ethical, legal, and governance issues. Societal impacts. 2024 Jun 1;3:100040.

[4] Ryan M. The social and ethical impacts of artificial intelligence in agriculture: mapping the agricultural AI literature. Ai & Society. 2023 Dec;38(6):2473-85.

[5] Ansari W. The Ethical Implications of Artificial Intelligence on the Human Society. Available at SSRN 4512301. 2023 Jul 16.

[6] Sanz JL, Zhu Y. Toward scalable artificial intelligence in finance. In2021 IEEE International Conference on Services Computing (SCC) 2021 Sep 5 (pp. 460-469). IEEE.

[7] Haefner N, Parida V, Gassmann O, Wincent J. Implementing and scaling artificial intelligence: A review, framework, and research agenda. Technological Forecasting and Social Change. 2023 Dec 1;197:122878.

[8] Sai S, Chamola V, Choo KK, Sikdar B, Rodrigues JJ. Confluence of blockchain and artificial intelligence technologies for secure and scalable healthcare solutions: A review. IEEE Internet of Things Journal. 2022 Dec 29;10(7):5873-97.

[9] Moro-Visconti R. Artificial Intelligence-Driven Digital Scalability and Growth Options. InArtificial Intelligence Valuation: The Impact on Automation, BioTech, ChatBots, FinTech, B2B2C, and Other Industries 2024 Jun 2 (pp. 131-204). Cham: Springer Nature Switzerland.
[10] Oikonomou EK, Khera R. Designing medical artificial intelligence systems for global use: focus on interoperability, scalability, and accessibility. Hellenic Journal of Cardiology. 2025 Jan 1;81:9-17.
[11] Sayed-Mouchaweh M, Sayed-Mouchaweh, James. Artificial Intelligence Techniques for a Scalable Energy Transition. Springer International Publishing; 2020.
[12] Govea J, Ocampo Edye E, Revelo-Tapia S, Villegas-Ch W. Optimization and scalability of educational platforms: Integration of artificial intelligence and cloud computing. Computers. 2023 Nov 1;12(11):223.
[13] Hammad A, Abu-Zaid R. Applications of AI in decentralized computing systems: harnessing artificial intelligence for enhanced scalability, efficiency, and autonomous decision-making in distributed architectures. Applied Research in Artificial Intelligence and Cloud Computing. 2024;7(6):161-87.
[14] Pazho AD, Neff C, Noghre GA, Ardabili BR, Yao S, Baharani M, Tabkhi H. Ancilia: Scalable intelligent video surveillance for the artificial intelligence of things. IEEE Internet of Things Journal. 2023 Mar 31;10(17):14940-51.
[15] Sakly H, Guetari R, Kraiem N, editors. Scalable Artificial Intelligence for Healthcare: Advancing AI Solutions for Global Health Challenges. CRC Press; 2025 May 6.
[16] Shivadekar S, Halem M, Yeah Y, Vibhute S. Edge AI cosmos blockchain distributed network for precise ablh detection. Multimedia tools and applications. 2024 Aug;83(27):69083-109.
[17] Bano S, Tonellotto N, Cassarà P, Gotta A. Artificial intelligence of things at the edge: Scalable and efficient distributed learning for massive scenarios. Computer Communications. 2023 May 1;205:45-57.
[18] Mishra A. Scalable AI and Design Patterns: Design, Develop, and Deploy Scalable AI Solutions. Springer Nature; 2024 Mar 11.
[19] Panda SP. Augmented and Virtual Reality in Intelligent Systems. Available at SSRN. 2021 Apr 16.
[20] Abisoye A, Akerele JI. A scalable and impactful model for harnessing artificial intelligence and cybersecurity to revolutionize workforce development and empower marginalized youth. International Journal of Multidisciplinary Research and Growth Evaluation. 2022 Jan;3(1):714-9.
[21] Raman R, Buddhi D, Lakhera G, Gupta Z, Joshi A, Saini D. An investigation on the role of artificial intelligence in scalable visual data analytics. In2023 International Conference on Artificial Intelligence and Smart Communication (AISC) 2023 Jan 27 (pp. 666-670). IEEE.
[22] Panda SP. The Evolution and Defense Against Social Engineering and Phishing Attacks. International Journal of Science and Research (IJSR). 2025 Jan 1.
[23] Newton C, Singleton J, Copland C, Kitchen S, Hudack J. Scalability in modeling and simulation systems for multi-agent, AI, and machine learning applications. InArtificial Intelligence and Machine Learning for Multi-Domain Operations Applications III 2021 Apr 12 (Vol. 11746, pp. 534-552). SPIE.

[24] Bestelmeyer BT, Marcillo G, McCord SE, Mirsky S, Moglen G, Neven LG, Peters D, Sohoulande C, Wakie T. Scaling up agricultural research with artificial intelligence. IT Professional. 2020 May 21;22(3):33-8.
[25] Meir Y, Sardi S, Hodassman S, Kisos K, Ben-Noam I, Goldental A, Kanter I. Power-law scaling to assist with key challenges in artificial intelligence. Scientific reports. 2020 Nov 12;10(1):19628.
[26] Shivadekar S, Kataria DB, Hundekar S, Wanjale K, Balpande VP, Suryawanshi R. Deep learning based image classification of lungs radiography for detecting covid-19 using a deep cnn and resnet 50. International Journal of Intelligent Systems and Applications in Engineering. 2023;11:241-50.
[27] Panda SP. Relational, NoSQL, and Artificial Intelligence-Integrated Database Architectures: Foundations, Cloud Platforms, and Regulatory-Compliant Systems. Deep Science Publishing; 2025 Jun 22.
[28] Shlezinger N, Ma M, Lavi O, Nguyen NT, Eldar YC, Juntti M. Artificial intelligence-empowered hybrid multiple-input/multiple-output beamforming: Learning to optimize for high-throughput scalable MIMO. IEEE Vehicular Technology Magazine. 2024 May 20;19(3):58-67.
[29] Samuel O, Javaid N, Alghamdi TA, Kumar N. Towards sustainable smart cities: A secure and scalable trading system for residential homes using blockchain and artificial intelligence. Sustainable Cities and Society. 2022 Jan 1;76:103371.
[30] Villegas-Ch W, Govea J, Gurierrez R, Mera-Navarrete A. Optimizing security in IoT ecosystems using hybrid artificial intelligence and blockchain models: a scalable and efficient approach for threat detection. IEEE Access. 2025 Jan 22.
[31] Mungoli N. Scalable, distributed AI frameworks: leveraging cloud computing for enhanced deep learning performance and efficiency. arXiv preprint arXiv:2304.13738. 2023 Apr 26.
[32] Panda SP. Artificial Intelligence Across Borders: Transforming Industries Through Intelligent Innovation. Deep Science Publishing; 2025 Jun 6.
[33] Cheetham AK, Seshadri R. Artificial intelligence driving materials discovery? perspective on the article: Scaling deep learning for materials discovery. Chemistry of Materials. 2024 Apr 8;36(8):3490-5.
[34] Panda SP, Muppala M, Koneti SB. The Contribution of AI in Climate Modeling and Sustainable Decision-Making. Available at SSRN 5283619. 2025 Jun 1.
[35] Shivadekar S. Artificial Intelligence for Cognitive Systems: Deep Learning, Neuro-symbolic Integration, and Human-Centric Intelligence. Deep Science Publishing; 2025 Jun 30.
[36] DeCost BL, Hattrick-Simpers JR, Trautt Z, Kusne AG, Campo E, Green ML. Scientific AI in materials science: a path to a sustainable and scalable paradigm. Machine learning: science and technology. 2020 Jul 14;1(3):033001.
[37] Klamma R, de Lange P, Neumann AT, Hensen B, Kravcik M, Wang X, Kuzilek J. Scaling mentoring support with distributed artificial intelligence. InInternational Conference on Intelligent Tutoring Systems 2020 Jun 3 (pp. 38-44). Cham: Springer International Publishing.
[38] Otaigbe I. Scaling up artificial intelligence to curb infectious diseases in Africa. Frontiers in Digital Health. 2022 Oct 21;4:1030427.

[39] Dasawat SS, Sharma S. Cyber security integration with smart new age sustainable startup business, risk management, automation and scaling system for entrepreneurs: An artificial intelligence approach. In2023 7th international conference on intelligent computing and control systems (ICICCS) 2023 May 17 (pp. 1357-1363). IEEE.
[40] Peteiro-Barral D, Guijarro-Berdiñas B. A study on the scalability of artificial neural networks training algorithms using multiple-criteria decision-making methods. InInternational Conference on Artificial Intelligence and Soft Computing 2013 Jun 9 (pp. 162-173). Berlin, Heidelberg: Springer Berlin Heidelberg.
[41] Kuguoglu BK, van der Voort H, Janssen M. The giant leap for smart cities: Scaling up smart city artificial intelligence of things (AIoT) initiatives. Sustainability. 2021 Nov 7;13(21):12295.
[42] Gowda D, Chaithra SM, Gujar SS, Shaikh SF, Ingole BS, Reddy NS. Scalable ai solutions for iot-based healthcare systems using cloud platforms. In2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) 2024 Oct 3 (pp. 156-162). IEEE.
[43] Awan MZ, Jadoon KK, Masood A. Scalable and effective artificial intelligence for multivariate radar environment. Engineering Applications of Artificial Intelligence. 2023 Oct 1;125:106680.
[44] Landin M. Artificial intelligence tools for scaling up of high shear wet granulation process. Journal of Pharmaceutical Sciences. 2017 Jan 1;106(1):273-7.
[45] Panda SP. Securing 5G Critical Interfaces: A Zero Trust Approach for Next-Generation Network Resilience. In2025 12th International Conference on Information Technology (ICIT) 2025 May 27 (pp. 141-146). IEEE.
[46] Mocanu DC, Mocanu E, Stone P, Nguyen PH, Gibescu M, Liotta A. Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. Nature communications. 2018 Jun 19;9(1):2383.
[47] Blanco L, Kukliński S, Zeydan E, Rezazadeh F, Chawla A, Zanzi L, Devoti F, Kolakowski R, Vlahodimitropoulou V, Chochliouros I, Bosneag AM. Ai-driven framework for scalable management of network slices. IEEE Communications Magazine. 2023 Nov 23;61(11):216-22.
[48] Sadek AH, Mostafa MK. Preparation of nano zero-valent aluminum for one-step removal of methylene blue from aqueous solutions: cost analysis for scaling-up and artificial intelligence. Applied Water Science. 2023 Feb;13(2):34.
[49] Cohen RY, Kovacheva VP. A methodology for a scalable, collaborative, and resource-efficient platform, MERLIN, to facilitate healthcare AI research. IEEE journal of biomedical and health informatics. 2023 Mar 20;27(6):3014-25.
[50] Adelodun AB, Ogundokun RO, Yekini AO, Awotunde JB, Timothy CC. Explainable artificial intelligence with scaling techniques to classify breast cancer images. InExplainable Machine Learning for Multimedia Based Healthcare Applications 2023 Sep 9 (pp. 99-137). Cham: Springer International Publishing.