AI Facilitates or Impedes Human Flourishing

Dr. Maria Lai-Ling Lam
LCC International University, Lithuania
Roundtable: AI and Machine Learning

Abstract

This paper builds on an extensive literature review of AI governance norms, reflecting on the power structure and injustices reinforced by AI, and decades of personal experience in healthcare and education. This paper applies Kanter and Bynum's (2022) flourishing ethics to highlight basic ethical principles for human flourishing in the context of AI. As a rapidly accelerating political, economic, cultural, and scientific force, AI is likely to exacerbate global inequality and poverty, impeding human flourishing due to the significant power imbalance between AI companies and emerging economies in less affluent countries. The paper raises critical questions to prevent amplifying these injustices and to reactivate core human values often overlooked by statistical optimization. It recommends implementing inclusive growth dialogues, developing "safe-AI" rather than "AI-safe" cultures, and steering AI-innovation toward labor-using rather than labor-saving for greater justice. Advanced economies should prioritize AI innovation and integration, providing a robust regulatory framework and capital tax to ensure these few AI companies compensate those whose lives are disrupted. Emerging markets and low-income countries should invest in digital infrastructure and a digitally competent workforce, adapting AI labor-using, rather than labor-saving purposes. The co-development process between society and technology should involve wider stakeholder participation to enhance human flourishing in an AI-driven world.

Keywords: flourishing ethics, labor-using AI, safe-AI, inclusive dialogues, global inequality

Introduction

Artificial Intelligence (AI) encompasses technologies that enable machines to sense, interpret, act, and learn to imitate human cognitive abilities (Manning, 2020). With the advent of Generative AI (GenAI), which uses sophisticated large language models to

rapidly generate new text and images from existing data, the technology has been widely adopted in our daily lives, bringing promises and risks (Cazzaniga et al., 2024; Gratton, 2024; IBM, 2023; Maslej et al., 2024). AI is supposed to serve our central ethical value: human flourishing (Kantar & Bynum, 2022). However, powerful nations, corporations, businesses, organizations, and individuals are seeking AI devices to help them achieve their narrow goals and disseminate their own interests and fostering a narrow sense of human flourishing in the process (VanderWeele, 2017). Without good AI governance and regulation, AI will exacerbate income and wealth inequality, education and health disparities, and increase social and political turmoil (Arnold, 2024; Cazzaniga et al. 2023; European Union, 2024; UNESCO, 2023; WHO, 2023). As AI evolves faster than the regulations, there is a pressing need for thoughtful and ongoing dialogues on how to guide its development, mitigate risks, and promote justice for human flourishing. Jack Clark and Ray Perrault, the co-directors of the Stanford Institute for Human-centered Artificial Intelligence, echo these opinions in their recent Artificial Intelligence Index Report (Maslej et al., 2023):

> AI will continue to improve and, as such, become a greater part of all our lives. Given the increased presence of this technology and its potential for massive disruption, we should all begin thinking more critically about how exactly we want AI to be developed and deployed. We should also ask questions about who is deploying it—as our analysis shows, AI is increasingly defined by the actions of a small set of private sector actors, rather than a broader range of societal actors.

This article is structured into four sections. The first section describes Kanter & Bynum's (2022) flourishing ethics and how AI can facilitate human flourishing. The second section describes how AI impedes human flourishing. The third section raises critical questions to prevent amplifying these injustices and to reactivate core human values often overlooked by statistical optimization. The fourth section recommends practices for the development and deployment of AI to enable human flourishing. The fifth section is the conclusion and points for future research areas.

Kanter and Bynum's Flourishing Ethics and AI facilitates Human Flourishing

Organizations and individuals are increasingly seeking AI devices to assist them as they try to quickly achieve their goals more efficiently at a lower cost, but this comes with growing risks as AI systems become more reliant on making decisions for us. Kantar & Bynum (2022) propose using the Flourishing Ethics approach to determine which ethical principles and values should be instilled in AI decision-making agents. They describe Flourishing Ethics as a set of related ethical theories that prioritize human flourishing as the central ethical value and with support by other ethical values (Kantar

& Bynum, 2022). They argue that humans share a common nature and require the following ethical values and principles for human flourishing: autonomy; supportive community; community with security, knowledge, opportunities and resources; justice; self-respect, and mutual respect. Kantar & Bynum (2022) wrote:

> Autonomy—the ability to make significant choices and carry them out—is a necessary condition for human flourishing. To flourish, people need to be part of a supportive community. The community should provide—as effectively as it can—security, knowledge, opportunities, and resources. To maximize flourishing within a community, justice must prevail. Respect—including mutual respect between persons—plays a significant role in creating and maintaining human flourishing….Self-respect also is important for human flourishing in order to preserve human dignity and minimize the harmful effects of shame, self-disappointment, and feelings of worthlessness (p. 603).

AI can automate and augment human capabilities to scale up effective and low-cost solutions quickly. This can facilitate human flourishing. For example, Singapore Eye Lesions Analyzer (SELENA) successfully implemented devices to automate the detection of referable diabetic retinopathy (DR) in Singaporean public hospitals with support from politicians, physicians, and patients. Politicians benefit from reduced healthcare costs and enhanced international reputations, physicians see reduced screening workloads, and patients receive results within one hour instead of weeks (Miller et al., 2024). SELENA's success in Singapore led to its adoption in African countries for DR grading. The DR devices support patients' autonomy by providing information in a way that helps patients make informed decisions and take early intervention about their health. Patients are supported by community with respect, security, and an abundance of resources. The DR also enables patients in African countries to be informed to take early intervention.

Another example is Med-PaLM 2 (2024), a large language model designed for the medical domain that can pass the medical exam, be a generalist, provide second opinions to patients, and deal with multiple imaging and disease.ChaptGPT4 generates original materials that are similar to the trained data, demonstrates symbolic reasoning, and shows a capacity for logic. Speech recognition devices can also help scientists to group similar chemicals together. Large data set from cities, such as the flow of automotive traffic and the footprint of passengers can be used to improve urban planning and city management (Williams, 2023).

Thus, AI may enable many people to access higher-quality education, health care, professional research, creative arts, and employment opportunities…etc (Askin et al., 2023; Bajwa et., 2021; Cukurova et al., 2023; Cukurova, 2024). New positions related to AI development skills, data science, computational modeling, ethics and more are being

generated to simulate and optimize current industry systems for greater competitive advantages. Research shows that high-skilled professionals who are exposed to AI will be shielded from being replaced because social norms and values prefer the existence of skilled professionals who are kept in the loop of the AI-augmented decision-making process (Cazzaniga et al., 2024). For example, the productivity and income of judges and doctors are expected to increase when they demonstrate the ability to integrate AI into their workflows for the benefit of their clients.

AI Impedes Human Flourishing

AI also has great potential for massive disruptions such as increased unemployment, out-of-control automation processes, and more. AI systems will exceed human capabilities as noted by Turing (1951): "Once the machine thinking method had started, it would not take long to outstrip our feeble powers." AI will affect our work, relationships, democratic systems, and national security. AI will exacerbate inequality and poverty around the globe (Aghion et al., 2021; Cazzaniga et al., 2024; Korinek et al., 2021). The gain of capital is much higher than the gain of labor wages given by AI as an intensive capital investment (Acemoglu, 2023; Acemoglu & Johnson, 2023, 2024; Acemoglu & Restrepo, 2022). Most of the benefits of growth will favor those at the top, resulting in wider income inequality in most countries. Many AI labor-saving devices serving the world, developed from a few AI "superstar companies," will result in lower actual GDP of emerging and low-income countries (Korinek et al., 2021). AI is predicted to have an unequal negative disruptive effect on people in different geographical and industry areas (Cazzaniga et al., 2024; Crawford, 2021; Goldman Sachs, 2024a, 2024b; McKinsey & Co., 2023). Powerful companies can easily shift their liabilities to humans involved in AI-augmented decision-making process (Tschider, 2024). Core human values such as mercy, dignity, and genuine human connections can easily be denied by AI based on probability statistics and the programming effort to optimize for particular groups' human preferences (Russell, 2022; Tasioulas, 2022).

AI can augment high-skilled professionals to have higher productivity but can also displace many people through automation. AI will likely worsen overall inequality and impede human flourishing for those most in need. IMF research finds that AI will endanger 33% of jobs in advanced economics, 24% in emerging economics, and 18 % in low-income countries (Alonso et al., 2020). AI will make a profound impact on the labor market quicker in advanced economics than in emerging economics and low-income countries because these developed countries' employment structures focus more on cognitive tasks (Cazzaniga et al., 2024). Some cities in developed countries that do not attract the investment of AI may face tremendous economic and political turmoil while more than 30% of their workers would face the risk of substitution by autonomous devices in a short period of time (Goldman Sachs, 2024a, 2024b; McKinsey & Co., 2023). The benefits of AI cannot be shared equitably given the current development and

deployment of AI in developed countries (Crawford, 2021). Although the developing and emerging economies seem to have less exposure to the negative impact of AI, their low-skilled labor can be easily substituted by labor-saving AI devices.

There is a significant power imbalanced between the select few AI companies that control the world market and emerging economies and low-income countries. A few AI companies located in a few powerful countries but serve the entire world economy. These "superstar" companies disproportionately located in high-income countries, do not bear the cost of employee displacement and welfare outside their own borders. They refine their AI models using data collected from emerging economies and low-income countries without compensation and provide advertisements or information to these regions at almost zero marginal cost. These few companies monopolize the international market and create high entry barriers for other firms, especially those in emerging economies and low-income groups. AI devices from these powerful companies will be easily and cheaply rolled out to developing countries and replace many unskilled workers in developing countries. These AI labor-saving devices will reduce investment in developing countries and their actual GDP (Alonso et al., 2020). Inequalities and poverty created by AI within developing countries will be greater than those in developed countries because developing countries often lack the institutional capacities to counteract harmful AI side effects or fully harness the benefits of AI (Cazzaniga et al., 2024). These few "superstar companies" enjoy winner-takes-all benefits, undermine the bargaining power of traditional labor, and do not pay any price for increasing social and political turmoil in countries with many unskilled workers and low AI readiness (Korinek et al., 2021).

When a small group of private technology companies control the AI foundation models, it can reinforce existing power structures and worsen labor conditions (Crawford, 2021; Lewchuk, 2017). Foundation models are trained on broad datasets that contain not only statistical/computational biases but also human and system biases. They may perpetuate human and system biases (Bommasani et al., 2023; Maslej et al., 2023; Sendak et al., 2020). Often, the data for training focuses on statistical/computational biases while ignoring historical human and systemic biases. For example, foundation models are frequently trained on data obtained from white people in developed countries, so that data is naturally embedded in their values. The complex layers of these algorithms are not explainable and transparent to users (Jin et al., 2024). For example, the defects of the foundation model are inherited by application models in the healthcare system, leaving users unable to understand how the AI system functions or fails. The foundation model can exacerbate social inequalities and fail to adapt to the multimodality of scanning equipment, new experimental technologies, or settings. Suppliers rarely report data quality assessments when presenting their model performance (Sendak et al., 2020). The Center for Research on Foundation Models must continuously urge developers to disclose data quality such that they can report

transparency index scores (Bommasani et al., 2024). There is no standardized data quality framework for similar AI-medical products (Rajesh et al., 2023; Sendak et al., 2020).

Many AI products in advanced economics need strategic alignment between organizational-administrative processes, AI technologies, and slack resources to do experimentation (Alami et al., 2024, Bommasani et al., 2023; Greenhalgh et al., 2017; Maslej et al., 2023, 2024). For example, hospital acquisition processes focus solely on the lowest price in the tender process, in the long run, they may end up paying more to update their software or risk using outdated software. It is also an inevitability for new technologies to lack compatibility with existing AI devices. Many AI products are not comparable and are operated across different hospitals or health care systems (Sendak et al., 2020). Generally, many AI devices present and utilized in healthcare originate from countries with extensive expertise and resources. When AI devices are presented as augmenting rather than replacing doctors, they gain more trust. In emerging economies and low-income countries with scarce resources and a shortage of doctors to support these devices, people may develop automation biases and trust these devices without much scrutiny, assuming comparable success in developed countries guarantees similar results in their own contexts. Later these countries have to pay higher prices for these AI medical devices (Rajesh et al., 2023).

AI may increase the productivity of low-skilled professionals as they can, in certain fields, quickly perform up to high-skilled professionals (Dell'Acqua et al., 2023). However, they may not enjoy higher income as the price of the products and services provided will lower given the increasing return of capital with reduced or the same labor. The illusion of mastery among low-skilled professionals may prevail. They may rely on AI and not adequately understand relevant functions or have the relevant knowledge when the AI tool is taken away or does not perform as expected (Felten et al., 2023). They will also lose opportunities to learn how to judge and become mature in their learning. Furthermore, AI also demoralizes those who cannot move to other industries while their jobs are automated. For example, AI disrupts these artists' livelihoods today. The mimicry model replaces these artists by using their often copyrighted art as training data and later creates new art imitating the original artist's creative properties or using the materials to create new properties without claiming any debt to the original art (Harris, 2023).

When the technocrats advocate for the advancement of AI that can imitate essential human characteristics like creativity, empathy, and learning, we may not respect those skills, characteristics, and virtues that take years to develop for humans. For example, empathy is an incarnate ability and needs to be developed through daily practices. If we think robots can provide us with emotional support, we will be hesitant to extend our efforts to express ourselves to other human beings since we can easily channel these

emotions to robots. This might solve a pressing temporal emotional problem but cannot develop our deeper connections to other human beings or to ourselves. People may lose their autonomy when they overly depend on AI tools for emotional support and decision-making (Sutton, 2024). It is unclear how prolonged reliance on AI for mental health support could impact clients' capabilities to connect with nature and others. Furthermore, the speed of spreading mass information that needs minimum intellectual and creative investment is shaping our habits of consuming information and our values of appreciating people's experience in skilled practices.

Given the commercial ambitions about investment in AI and the AI national race, the public may defer ethically and politically contentious AI issues to technocratic and government elites. Furthermore, we humans are divided among nations and these nations are often tempted to win the AI national race at the expense of basic human rights and values. With the current geopolitical context and space race-like event to develop the technology, it may become possible to mistakenly rationalize the value of mass killing, deportation, or even genocide for national protection. Some people perceive fear when their own countries constrain the capabilities of AI while their enemies fully utilize the capabilities of AI. Using AI for national surveillance can be justified at the expense of individual privacy or human life.

Key AI Questions and Core Human Values

AI generally focuses on the economics of scale, economic prosperity, and maximizing some notions of expected utility. The value of optimization of human preference may deny core basic human values (Tasioulas, 2022). AI is changing our understanding of the good life, the virtues of patience, thinking, knowledge, consciousness, emotions, society, good and evil, and the ultimate nature of the universe. AI is evolving rapidly, creating many potential promises and harms. AI can be deployed on a large scale and at a speedy rate without knowing whether the AI has been misaligned with current social norms and human values (Hinton, 2023a, 2023b; Russell, 2019, 2022, 2024). Many unknown emergent patterns will be generated from large language models while people have to keep fine-tuning preexisting data sets to increase their accuracy and fairness (Lapata, 2023; Woldridge, 2023; Zewe, 2022). When AI shares thousands of copies of the same model on a particular section of a large data set, human experts cannot as effectively and easily share their findings from different disciplines as AI models do. It must also be considered that eventually AI could escape human control and become an unaccountable machine (Davies, 2024). Wiener (1960) has already warned us that human can easily over-rely on automated machines to make decisions for themselves without questioning the purpose and potential negative consequences of using the automated machine. He said,

> Though machines are theoretically subject to human criticism, such criticism may be ineffective until long after it is relevant. To be effective in warding off disastrous consequences, our understanding of our man-made machines should in general develop at the same rate with the performance of the machine. By the very slowness of our human actions, our effective control of the machines may be nullified. By the time we are able to react to the information conveyed by our senses and stop the car we are driving, it may already have run head on into a wall (p. 1355).

The overarching ethical principle in the AI system is adopting preference-based utilitarianism. AI systems are designed to maximize the fulfillment of human preferences that are shaped by human experience and cannot be easily known and quantified (Russell, 2022; Tasioulas, 2022). Societal ethics are unfortunately reduced to precision and an optimization of human preferences in the AI system. Furthermore, these preferences are shaped by powerful economic and government actors. They shape the preferences for their own power, not social wealth. Will the utilitarian thinking in the AI system confine us and neglect other human values such as justice, fairness, charity, hope, autonomy, friendship and play?

Bommasani et al. (2021) remind us of the unknown emerging properties of the foundation model which requires deep interdisciplinary collaboration. Defects of the foundation models will be inherited by specific applications and also incentivize harmonization. They said,

> Despite the impending widespread deployment of foundation models, we currently lack a clear understanding of how they work, when they fail, and what they are even capable of due to their emergent properties.

AI is trained on pre-chosen sets of data for specified tasks. They cannot handle situations not captured by the pre-trained data set. Sheikh et al. (2023) said, "AI is associated with a distinct techno-economic paradigm characterized by continuous change to products and services, a largely hidden vertical structure of hardware and software, and the potential for technology to act autonomously" (p. 96). Using statistical optimization to find the patterns in the data may result in missing many interesting phenomena and alternatives as they are not described by the bell curve in the statistical analysis. Data used to train the foundation model often inherit all those historical and systemic biases, which can be transferred to the application models and further marginalize many minority groups. The Center for Research on Foundation Models remains dissatisfied with developers' data disclosure behavior (Bommasani et al., 2024). The explainability of the current foundational model and an assessment of the quality of data in the training of the foundation model are both still lacking. These foundation models can exacerbate social inequalities. It needs countless investments in training the

data and refining the foundation model to correct mistakes such as historical biases and hallucinations.

AI challenges all industries to reflect on the meaning of their existence. For example, journalists are encouraged to know how to use AI data to adapt to changing consumers' reading habits and be faithful to the mission to serve the civic information space. They will keep public trust and institutional integrity by informing the public of any misinformation of the data (Murray, 2024). Educators must reconsider the purpose of education and adapt their pedagogies to enhance access and equity, foster human connections, enrich learning experiences, and build emotional resilience. Students should also be encouraged to apply compassion and creativity to tackle real-world challenges.

We need to be aware of the new paradigm created by AI systems and the societal and communal core values that might be denied or ignored by AI. When AI issues are treated as primarily technical issues solved by experts, we may ignore our own character development and other people's well-being. What kind of person will we become when we practice to achieve certain outcomes at faster rates and lower costs? What virtues will be cultivated when patience is not valued? When corporate leaders are anxious to seek opportunities created by AI and want to get ahead of their competition, they have shown a willingness to take higher risks to speed up and scale their deployment of AI that may end up harming human beings and society in the long-run (Murray, 2024). Will safety be treated only after problems arise? Who is most vulnerable to the risks of faulty AI systems in the current global market economy? When we are used to communicating with AI machines, will we forget that it is a machine? Will our communication with AI machines increase our illusion of mastery and reduce our implementation abilities in reality? Will the loser in this ruthless AI competition be justified to bear the cost of AI disruption or AI mistakes created by emerging properties of generative AI systems? Will our rationalization of AI adaptability and responsibility deny basic human values like justice and dignity?

We need to think deeply about how we can have a broader engagement of people in the development and deployment of AI in different sectors such that AI will not "amplify and reproduce the forms of power it has been deployed to optimize" (Crawford, 2021, p. 224). How should flourishing ethics be incorporated into the AI system? How should our education system equip our students to know how to work closely with AI and also how to develop good practices that better incorporate broader values for human flourishing?

We need to demand AI companies to invest in the development of safe AI first and research human interfaces in AI design. What values are advocated in the process of development and deployment of AI systems? Can humans control AI systems? Do we

expect scientific and bureaucratic expertise to devise efficient and effective mechanisms for securing the goals of AI regulations even though AI is rapidly evolving beyond the benchmarks set by industries and politicians? Now, many business enterprises are motivated by the short-term benefits of integrating AI into their workplace and are doing their best to adapt as quickly as possible. We perceive that we can control AI when we are allowed to interrupt AI systems (Baker, 2020).

Why do we value mercy from other humans and not from robots? Why will patients be angry when they find out that the apology letters they received from doctors were written by AI? Why do humans still value human connections in many crucial legal and medical judgments and decisions? Tasioulas (2022) explained we need human judges rather than robot judges because we value human accountability while "autonomous machines that do not have a share in human solidarity and cannot be held accountable for their decisions in the way that a human judge can "(p. 238). We want "the ideal of reciprocity among fellow citizens that is central to the rule of law" (p. 237) even though the autonomous machine may have better transparency, procedural fairness, and explainability than human judges.

What is the intrinsic value of actual labor? Will the benefit of automation convince employers to replace select occupations with AI? What is the collective vision of an AI-ecosystem that supports human flourishing? Tasioulas (2022) concluded that "our focus must be properly integrating AI technology into a culture that respects and advances the dignity and well-being of humans, and the nonhuman animals with whom we share the world, rather than on the highly speculative endeavor of integrating the dignity of intelligent machines into our existing ethical framework" (p. 240). How do we preserve our genuine human interactions in an increasingly automated work environment?

Can humans cede basic levels of thinking and skill-acquisition processes to AI and instead focus on higher-order thinking skills based on AI outputs? The concern is that when basic skills are ceded to machines, humans lose the opportunity to develop higher-ordered skills through scaffolding. This could result in many superficial users who over-rely on machines and accept AI solutions without questions. Conversely, sophisticated users will know how to effectively prompt and input tasks for AI, using their prior knowledge to judge the validity of the information. Some may apply ethical principles when working with AI, continually questioning their purposes for using AI and resisting being controlled by machines. Their experience of using AI as professional development tools and their communication with AI must be learned by superficial users.

AI co-evolves with society and changes our expectations about the functions of AI and the meanings of intelligence provided by AI. As AI becomes associated with prolonged processes of social and technological co-evolution, we can collectively question how

economic and political structure, power, and the dynamics of social (in) justices created by AI generate more dialogues on what needs to be changed for more people to economically and socially flourish (Fassin & Das, 2021). Without productive questioning, we cannot easily adapt to the changes and expect a better return without paying attention to the direction of the technology. Sheikh et al. (2023b), "The development of system technologies is often unpredictable, and their effects cannot be fully anticipated…The generic nature of such technologies implies that they have the potential to affect all public values" (p. 100). It is essential to provide a better environment for the development and deployment of AI for public benefit and develop a collective vision of better ecosystems for human flourishing.

Recommended Practices

Both the World Health Organization (WHO, 2023) and the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2023) affirm the immense potential of AI to transform healthcare and education. They propose to develop safe and equitable AI ecosystems to promote human flourishing. For example, AI should enhance human capabilities to build inclusive digital futures for all and should never be an authoritative source of knowledge on whatever topic it engages with (UNESCO, 2023). In health care policy, AI should protect autonomy; promote well-being, human safety, and the public interests, ensure transparency, explainability, and intelligibility; foster responsibility and accountability; ensure inclusiveness and equity; and promote AI that is responsive and sustainable (WHO, 2023). Inclusive growth dialogues such as global tax regimes for the digital age, global competition policy, intellectual property rights, data and information policy, and new transparency regulations must be adopted to share economic prosperity around the globe (Geiger & Iaia, 2024; Lee, 2024). AI innovation should be steered toward more labor-using and greater shared prosperity (Korinek et al., 2021; Partnership on AI, 2022). Advanced economic should prioritize AI innovation and integration for better products and services for human flourishing, instead of just focusing on saving labor and lowering the cost of production (Aghion et al., 2021). They should provide robust regulatory frameworks to guide these "superstar" AI companies. Emerging economies and lower-income countries are supported to have more investment in digital infrastructure and a digitally competent workforce. Emerging economies and poor countries have to discern what AI-labor devices, rather than AI-labor-saving devices should be adopted.

The public is educated enough to know what AI is and is incentivized enough to balance the dominant perspectives of a select few rich technology companies. "An efficient and effective government, a well-functioning financial system, the absence of corruption, and civic stability are all important in supporting families, work, education, and religious communities in the promotion of individual flourishing; and the study of how more macro-and state-level factors influencing human flourishing is needed as

well." (VanderWeele, 2017). Russell (2024) advised that the only visible long-term option is to have mathematically guaranteed and provably beneficial AI. He urged the development of safe AI rather than an AI-safe culture. We need to construct and design safe AI first and determine the goals first before the design. This recommended practice goes against the current norms of the development and deployment of AI that seek to develop AI capabilities first, then constrain the capabilities for safety, and finally define the goals of the AI after mistakes are found in the pre-existing training data.

Policy-makers and regulators are recommended to introduce all technical, social, economic, and scientific dimensions of AI systems in society (Bommansani et al., 2023) through the following five key tasks: clarify what AI is and focus on actual risks and opportunities (i.e., demystification); create a functional ecosystem to make AI work (i.e., contextualization); involve diverse stakeholders from civil society to address relevant values and interests affected by the use of AI technology (i.e., engagement); develop a directive framework (i.e., regulation); and engage wisely with other global actors (i.e., societies) (Sheikh et al., 2023b). These five tasks steer the process of co-development between technology and society. The progress of AI and the shared benefits of AI are most likely to be maximized in democratic societies that allow many stakeholders to contribute and control the direction of AI's development (Acemoglu & Johnson, 2023). Thus, human beings need to be conscious of the dynamics between AI and the social, economic, political, and historical context that supports AI.

With good governance, AI can facilitate human flourishing. The good governance of AI encompasses a broad array of "processes of governing, whether undertaken by a government, market, or network, whether over a family, tribe, formal or informal organization, or territory, and whether through laws, norms, power, or language" (Bevir, 2012, p. 1). Individual developers must integrate ethical and legal issues much earlier in their design stage of AI. Organizational users must implement quality assurance and data governance policies. Local and international governments must enforce modified copyright laws, national guidelines, and intellectual property laws in an AI-driven world. We all need to address the three most important ethical issues raised by AI: privacy and surveillance, bias and discrimination, and accountability and transparency. The meaning of privacy indeed varies in different contexts, influenced by cultural, legal, and social factors. Understanding these nuances is critical for developing AI systems that respect privacy across diverse settings.

Conclusion and Future Research

AI is inevitably changing our work, values, and lives. With better public education about what AI is and a broad engagement of the public about the direction of AI, we can rebalance the power between AI technology companies and consumers. We must know what AI is and understand its limitations. The effectiveness of AI depends on the

contexts that support AI. We need to be conscious of our core human values such as genuine human contact and the dignity of work that are gradually being replaced by AI as the development of technology mainly focuses on the economics of scale, efficiency, and lower costs. The adoption process of mature AI technology into the familiar workflow needs to be monitored. We need safe AI rather than an AI safe culture! AI-innovation should be steered toward labor-using rather than labor-saving for greater justice. Advanced economics should prioritize AI innovation and integration, providing a robust regulatory framework and capital tax to ensure these few AI companies compensate those whose lives are disrupted. Emerging markets and low-income countries should invest in digital infrastructure and a digitally competent workforce, adapting AI labor-using, rather than labor-saving purposes. The co-development process between society and technology should involve wider stakeholder participation to enhance human flourishing in an AI-driven world. More stakeholders need to be invested in the discourse around AI systems and demand higher safety and security standards from these technology companies. There is no autonomous system that must be held accountable for the result: humans and human lives will always be held accountable. Inclusive global dialogues that enable emerging economies and low-income countries to have voices and be compensated fairly in the global competition policy, intellectual property rights, data and information policy, and new transparency regulations must be adopted.

More research needs to be conducted about human-system interaction and communication (Eshelman et al., 2012; Lyreskog et al., 2023; Vaill, 1989).

These are future research questions:

1.   How to solve AI's inequality problems?
2.   How can the benefits of AI be shared by more people, specifically those on the margins of power and wealth?
3.   What new skills and knowledge do humans need to effectively communicate with AI, especially when steering AI innovation for labor-using rather than labor-saving for greater justice?
4.   How do the development and deployment of AI challenge existing legal norms, such as copyright protection and antitrust laws, given that AI foundation models are currently being monopolized by a few companies?
5.   What are the challenges posed by fragmented legal regulations governing AI when different countries do not harmonize their AI regulations while competing for national security interests?

Our current choices in understanding, using, and regulating AI systems will determine whether AI tools enable human flourishing or not. The integration of AI into our current social-political-technology systems must address current challenges such as

data quality, technical infrastructure, organizational capacity, responsible practices, global inequality and poverty, and power imbalance between the few AI "superstar" companies and emerging economies and low-income countries. Without proper interventions and steering AI innovation toward labor-using practices and greater shared prosperity, social and wealth inequality could increase, even among knowledgeable workers in developed countries. AI will increase inequality, poverty, and social and political unrest around the globe when inclusive growth dialogues are not quickly implemented. More people need to be educated about what AI is and collectively demand better AI ecosystems that respect the dignity of work, autonomy, justice, and community development.

Acknowledgments

About the Author

Maria Lai-ling Lam, PhD, is currently the Chair of the International Business Administration Department at LCC International University in Klaipėda, Lithuania. She holds a Ph.D. in Marketing and Organizational Behavior from George Washington University, an M.A. in Religious Studies, and both an M.B.A. and B.B.A. from the Chinese University of Hong Kong.

Maria has taught various marketing, organizational behavior, and strategy courses to undergraduate and graduate students at four Christian universities in the United States for nineteen years and at several universities in Hong Kong for eight years.

She has conducted numerous seminars and workshops at prominent universities in China and Japan. Maria has published one book and over 100 peer-reviewed articles and book chapters. Her intellectual contributions cover areas such as U.S.-China business trust relationships, corporate social responsibility, empathy, human flourishing, management education, cybersecurity, smart cities, diversity, equity, and inclusion.

Correspondence concerning this article should be Emailed to: mlam@lcc.lt

References

Alami, H., Lehoux, P., Papoutsi, C. *et al*. (2024). Understanding the integration of artificial intelligence in healthcare organisations and systems through the NASSS

framework: a qualitative study in a leading Canadian academic centre. *BMC Health Serv Res* 24, 701. https://doi.org/10.1186/s12913-024-11112-x

Acemoglu, D. (2023). Harms of AI. In J.B Bullock, Y-C Chen, J. Himmelreich, V. M. Hudson, A. Korinek, M. M. Young, & B. Zhang (Eds.), *The Oxford Handbook of AI Governance* (1st ed., pp. 660-760). Oxford University Press

Acemoglu, D., & Johnson, S. (2023). *Power and progress: our thousand-year struggle over technology and prosperity*. PublicAffairs.

Acemoglu, D., & Johnson, S. (2024). *Learning from Ricardo and Thompson: Machinery and labor in the early industrial revolution, and in the age of AI*. MIT Shape the Future of Work.

Acemoglu, D., & Restrepo, P. (2022). Tasks, Automation, and the Rise in U.S. Wage Inequality, *Econometrica*, 90 (5), 1973–2016

Aghion, P., Antonin, C., & Bunel, S. (2021). Technological progress, artificial intelligence, and inclusive growth. *IMF Working Papers*, 2021(123), 1-45. https://www.imf.org/en/Publications/WP/Issues/2021/06/11/Technological-Progress-Artificial-Intelligence-and-Inclusive-Growth-460695

Alonso, C., Berg, A., Kothari, S., Papageorgiou, C., & Rehman, S. (2020). *Will the AI revolution cause a great divergence*? IMF Working Paper No. 20/184. International Monetary Fund. https://doi.org/10.5089/9781513556505.001

Arnold, M. (2024 June 26). Central banks urged to 'raise game' on use of artificial intelligence. *Financial Times*.

Askin, S., Burkhalter, D., Calado, G., & El Dakrouni, S. (2023). Artificial Intelligence Applied to clinical trials: opportunities and challenges. *Health and technology*, *13*(2), 203–213. https://doi.org/10.1007/s12553-023-00738-2

Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future Health Care Journal*,8(2), e188-e194. doi: 10.7861/fhj.2021-0095. PMID: 34286183; PMCID: PMC8285156.

Baker, S. (2020). *The Boeing 737 Max crashes have revived decades-old fears about what happens when airplane computers become more powerful than pilots*. Retrieved March 30, 2024 from Business Insider

Bevir, M. (2012). *Governance: A very short introduction*. OUP, Oxford.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, V., Cardie, C., Catanzaro, B., Chiang, M., Clark, J., De Sa, C., Demszky, D., Eismann, S., ... & Liang, P. (2021). *On the opportunities and risks of foundation models*. Center for Research on Foundation Models (CRFM), Stanford Institute for Human-Centered Artificial Intelligence (HAI), Stanford University. arXiv:2108.07258v3.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, V., Cardie, C., Catanzaro, B., Chiang, M., Clark, J., De Sa, C., Demszky, D., Eismann, S., & Liang, P. (2023). *On the opportunities and risks of foundation models*. Center for Research on Foundation Models (CRFM), Stanford Institute for Human-Centered Artificial Intelligence (HAI), Stanford University. arXiv:2108.07258v3.

Bommasani, R., Klyman, K., Longpre, S., Kapoor, S., Maslej, N., Xiong, B., Zhang, D., & Liang, P. (2024). *The foundation model transparency index*. Center for Research on Foundation Models, Stanford University. https://crfm.stanford.edu/fmti/May-2024/index.html

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E. J., & Tavares, M. M. (2024). *Gen-AI: Artificial intelligence and the future of work*. IMF Staff Discussion Notes, 2024 (001), International Monetary Fund, Washington, DC. https://doi.org/10.5089/9798400262548.006

Crawford, K. (2021). Conclusion: Power. In *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence* (pp. 211–227). Yale University Press. https://doi.org/10.2307/j.ctv1ghv45t.10

Cukurova, M., Miao, X., & Brooker, R. (2023). Adoption of Artificial Intelligence in Schools: Unveiling Factors Influencing Teachers' Engagement. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O.C. Santos, & V. Dimitrova (Eds.), *Artificial Intelligence in Education* (pp. 151–163). Lecture Notes in Computer Science (Vol. 13916). Springer, Cham

Cukurova, M. (2024). The interplay of learning, analytics, and artificial intelligence in education. *arXiv*

Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *Harvard Business School Working Paper 24-013, 1-58*

Davies, D. (2024). *The unaccountability machine: Why big systems make terrible decisions-and how the world lost its mind*. London: Profile Books Ltd.

Eshelman, G. L.; Lam, M.L.L.; and Cook, M. J. B. (2012). Three contributing factors to effective utilization of technology in management education and practice: Personhood, mindfulness, and meditation. *Journal of the North American Management Society*, 6 (1), Article 4.

European Union (2024 May 21) *Artificial intelligence (AI) act: Council gives final green light to the first worldwide rules on AI -* Consilium (europa.eu)

Felten, E. W., Raj, M., & Seamans, R. (2023). *How will Language Modelers like ChatGPT Affect Occupations and Industries?* SSRN. Available at https://ssrn.com/abstract=4375268 or http://dx.doi.org/10.2139/ssrn.4375268

Geiger, C., & Iaia, V. (2024). *The forgotten creator: Towards a statutory remuneration right for machine learning of generative AI. Computer Law & Security Review*, 52, 105925.

Goldman Sachs. (2024a). *The US labor market is automating and becoming more flexible.* Retrieved April 25, 2024, from goldmansachs.com

Goldman Sachs. (2024b). *Gen AI: too much spend, too little benefit?* Retrieved June 27, 2024, from goldmansachs.com

Gratton, L. (2024). *What AI means for human capital.* Retrieved March 4, 2024, from https://sloanreview.mit.edu/video/what-ai-means-for-human-capital.

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). *Beyond Adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies.* Journal of Medical Internet Research, 19(11), e367. https://doi.org/10.2196/jmir.8775

Harris, G. (2023, February 28). '*AI will become the new normal': How the art world's technological boom is changing the industry.* Retrieved August 15, 2025, from theartnewspaper.com

Hinton, G. (2023a). "Godfather of AI" Geoffrey Hinton Warns of the "Existential Threat" of AI | Amanpour and Company. Retrieved March 4, 2024 from youtube.com

Hinton, G. (2023b). Prof. Geoffrey Hinton - "Will digital intelligence replace biological intelligence?" Romanes Lecture. Retrieved March 4, 2024 from youtube.com

IBM. (2023). *The augmented workforce for an automated, AI-driven world*. IBM Institute for Business Value (IBV).

Jin, Q., Chen, F., Zhou, Y., Smith, J., Johnson, A., Brown, L., Davis, M., Garcia, R., Martinez, P., Wilson, T., Anderson, K., Thomas, H., Jackson, W., Harris, G., Martin, B., Lee, C., Perez, S., Clark, N., & Young, E. (2024). Hidden flaws behind expert-level accuracy of multimodal GPT-4 vision in medicine. *npj Digital Medicine, 7*, 190. https://doi.org/10.1038/s41746-024-01185-7

Kantar, N., & Bynum, T. W. (2022). Flourishing ethics and identifying ethical values to instill into artificially intelligent agents. *Metaphilosophy*, 53:599-604. Metaphilosophy LLC and John Wiley & Sons Ltd.

Korinek, A., Schindler, M., & Stiglitz, J. E. (2021). *Technological progress, artificial intelligence, and inclusive growth* (IMF Working Paper No. 2021/166). International Monetary Fund. https://doi.org/10.5089/9781513583280.001

Lapata, M. (2023, October 25). What is generative AI and how does it works. *Digital Habitats*.

Lee, E. (2024). Promoting progress: Authorship in the age of AI. *Florida Law Review*, Vol. 76. Available at SSRN or via DOI: 10.2139/ssrn.4609687.

Lewchuk, W. (2017). Precarious jobs: Where are they, and how do they affect well-being? *The Economic and Labour Relations Review*, 28(3), 402-419. https://doi.org/10.1177/1035304617722943

Lyreskog, D. M., Zohny, H., Singh, I., & Savulescu, J. (2023). The ethics of thinking with machines: Brain-computer interfaces in the era of artificial intelligence. *International Journal of Chinese & Comparative Philosophy of Medicine.*

Manning, C. (2020). *Artificial intelligence definitions*. Stanford University. Retrieved from https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf1

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). *Artificial intelligence index report 2023*. Stanford Institute for Human-Centered Artificial Intelligence. https://aiindex.stanford.edu/report/2023

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). *Artificial intelligence index report 2024*. Stanford Institute for Human-Centered Artificial Intelligence. https://aiindex.stanford.edu/report/2024

McKinsey & Co. (2023). The-state-of-ai-in-2023-generative-ais-breakout-year_vf.pdf (mckinsey.com)

Med-PaLM 2 (2024). Med-PaLM: A Medical Large Language Model - Google Research

Miller, S., Gomulya, D., & Rao-Kachroo, M. (2024). EyRIS: From the lab to the market. *Asian Management Insights*. 9_AMI_Mar24_Eyris.pdf (smu.edu.sg)

Partnership on AI (2022). Redesigning AI for shared prosperity: An agenda. Technical Report. Available at Shared Prosperity Guidelines - Partnership on AI

Rajesh, A. E., Davidson, O. Q., Lee, C. S., & Lee, A. Y. (2023). Artificial Intelligence and Diabetic Retinopathy: AI Framework, Prospective Studies, Head-to-head Validation, and Cost-effectiveness. *Diabetes care*, 46(10), 1728–1739. https://doi.org/10.2337/dci23-0032

Russell, S. (2019). *Human Compatible: AI and the Problem of Control.* London: Allen Lane.

Russell, S. (2022). Artificial Intelligence and the Problem of Control. In: Werthner, H., Prem, E., Lee, E.A., Ghezzi, C. (eds) *Perspectives on Digital Humanism*. Springer, Cham. https://doi.org/10.1007/978-3-030-86144-5_3

Russell, S. (2024, April 25). *AI: What if we succeed*. YouTube.

Sendak, M. P., D'Arcy, J. Kashyap, S., Gao, M., Nichols, M., Corey, K., Ratliff, W., & Balu, S. (2020). A path for translation of machine learning products into health care delivery. *European Medical Journal Innovations*. https://doi.org/10.33590/emjinnov/19-00172

Sheikh, H., Prins, C., & Schrijvers, E. (2023a). Artificial Intelligence: Definition and Background. In: *Mission AI. Research for Policy*. Springer, Cham.

Sheikh, H., Prins, C., & Schrijvers, E. (2023b). AI as a System Technology. In: *Mission AI. Research for Policy*. Springer, Cham.

Sutton, J. (2024). *Revolutionizing AI therapy: The impact on mental health care*. Retrieved June 8, 2024 from Positive Psychology: https://www.positivepsychology.com.

Tasioulas, J. (2022). Artificial Intelligence, Humanistic Ethics. *Daedalus*; 151 (2): 232–243.

Tschider, C. (2024). Humans Outside the Loop. 26 *Yale J.L. & Tech*. 324 (2024), Available at SSRN: https://ssrn.com/abstract=4580744 or http://dx.doi.org/10.2139/ssrn.4580744

Turing, A. M. (1951). Can digital machines think? *Radio broadcast, BBC Third Programme.* Typescript available at turingarchive.org.

UNESCO (2023). *Guidance for generative AI in education and research.*

VanderWeele, T. J. (2017). On the promotion of human flourishing. *PNAS*, 8148-8156, Aug. 1, 114 (31).

Vaill, P. (1989). *Managing as a performing art: New ideas for a world of chaotic change.* San Francisco: Jossey-Bass Inc., Publishers.

WHO (2023*). Ethics and governance of artificial intelligence for health.*

Wiener, N. (1960). Some moral and technical consequences of automation. *Science* 131:1355-58.

Williams, S. (2023). IAAC Lecture Series – Data Action: Using Data for a Public Good. YouTube

Woldridge, M. (2023). *What's the future of generative AI* [Video]. YouTube.

Zewe, A. (2022). A technique to improve both fairness and accuracy in artificial intelligence | MIT News | Massachusetts Institute of Technology