Fulbright Chronicles

Volume 4, Issue 1: (November 2025)

In this Issue:

Artificial Intelligence • Education • Equity • Bias • AI-Generated • Text Detection • Governance • Human-centered AI • Service-Learning • Information Diplomacy • Infodemic Management • Experiential Education • AI Literacy • Digital Media Literacy • AI in K-12 Education • Digital Divide • Ethics • Innovation. • Africa • eLearning • Reality • Fulbright • Thailand • Cognitive Neuroscience • Interdisciplinary Collaboration • Creativity • Authorship • Intellectual Property • Digital Transformation • Generative AI (GenAI) • Information Literacy • Critical Thinking • Digital Pedagogy • Academic Integrity •

Fulbright Chronicles is an independent not for profit 501 (c) (3), open access, peer reviewed, quarterly periodical with contributions by and for the global Fulbright community. The journal is overseen by a global Editorial Board.

ISSN: 2998-4998

About and Contributing

The journal provides a welcoming space for Fulbrighters to share their work and reflections on global issues with a broad audience. It features thoughtful, accessible articles that reflect on how Fulbright experiences have contributed to knowledge and cross-cultural understanding, or that comment on contemporary issues that affect the Fulbright program or cultural and educational exchange more broadly.

The *Fulbright Chronicles* can only succeed with the engagement of the Fulbright community. The editors strongly encourage Fulbrighters to contribute articles or commentaries on topics related to your research and practice and the critical issues of our times. Author Guidelines are available on our website (www.fulbright-chronicles.com).

The journal is an independent publication, overseen by the Editorial Board members. For further information, visit the *Fulbright Chronicles* site (www.fulbright-chronicles.com).

Table of Contents

A New Era for Fulbright Chronicles

Bruce B. Svare, Habiba I. Atta, Melanie C. Brooks, and Kevin F. F. Quigley

As we wrote in our last issue, this third year of the *Fulbright Chronicles* would be a year of major milestones and significant changes. This issue marks another important transition. We announce with great enthusiasm that we are now a completely independent 501(c)(3) not-for-profit organization registered in the United States as *The Fulbright Chronicles Foundation*.

We are grateful for the previous in-kind support we received from the United Kingdom Fulbright Commission, but it was time to spread our wings and strike out on our own. At the forefront of this decision was our desire to more fully represent the scholarly activities and important work of Fulbright alumni from around the world, and to do so without any possible restraint that would compromise freedom of expression.

Along with our new independent status is the responsibility to our readers to press forward with our mission to focus on the impact of the incredible work that Fulbright alumni do around the world. This independent status lets the *Chronicles* play a greater role in advocacy for the Fulbright Program itself. We will perform both roles responsibly and with integrity.

As our readers know, these are perilous times for international exchange programs, and we need to fight for their survival. We must also actively support efforts that increase our understanding of important threats to our world including lack of education, discrimination, extreme inequality, demographic explosion, climate change, mass migration, violence, and poverty, among other issues. These times demand that Fulbrighters continue their important work to engage in strategic and culturally sensitive scholarship that is done collaboratively and addresses important global issues affecting all of humankind.

As an independent journal, we are reliant upon donations to publish this journal. We are an all-volunteer, internationally represented, Fulbright alumni group that is committed to giving back to the Fulbright exchange program. None of us are paid; we need funds to handle the expenses involved with hosting, archiving, creating and loading content to our website.

If we are to survive as an independent journal that is freely accessible and does not use subscription or submission fees, page charges or a paywall, we ask that you help underwrite our costs. Please let us know that you value the mission of *Fulbright Chronicles* and the independent, professional product that is published. You can seamlessly donate on our website (www.fulbright-chronicles.com). We will be grateful for any level of support that you can provide.

In this period of transition for *Fulbright Chronicles*, it is fitting that the present edition of the journal is a themed issue addressing artificial intelligence (AI) and its impact upon education and the social sciences. Masterfully edited by two of our associate editors, Jose Caetano and Polat Goktas, this issue is a diverse collection of commentaries from Fulbright alumni representing many different disciplines and regions of the world. The topic is one that is critically important for just about everything we do. Our readers will find their commentaries to be insightful and stimulating as we try to navigate the relentless impact of AI on our lives. Because AI is such a new topic for everyone including Fulbright alumni, you will note that these articles are less about our stated goal of examining/representing the impact of the Fulbright experience and more about the reflections of the authors on where the field of AI is going in the future. Please read the introduction to this issue by Jose and Polat before diving into the individual articles.

"If we are to survive as an independent journal that is freely accessible and does not use subscription or submission fees, page charges or a paywall, we ask that you help underwrite our costs."

This issue also provides book reviews regarding various topics in AI. The books, authored by Fulbrighters and reviewed by Fulbright alumni, are a testament to the impact of AI on our world today.

As we look forward to the "new" era of *Fulbright Chronicles*, we know that we can depend upon our readers for input and support. Together, we can continue the mission of the journal to present the best of the Fulbright Program and the scholarship and leadership that its alumni consistently advance. Please enjoy this special issue and stay connected to Fulbright Chronicles!

Beyond Code and Curriculum: Reimagining Education Through AI and Social Science Lenses

Jose Caetano and Polat Gotkas

Artificial intelligence (AI) is reshaping education. And it is here to stay. In classrooms, from elementary schools to universities, algorithms are increasingly making decisions about what, how, and when we learn. Yet, the debate about AI in education has too often been polarized by technical solutions, overlooking the social, cultural and ethical dimensions that equally shape its impact. This special issue (SI) of the Fulbright Chronicles brings together Fulbright alumni from many different fields to imagine a future where AI in education is not only efficient but also fair, inclusive, and culturally responsive.

This SI was motivated by our belief that technological progress alone cannot guarantee better education for all. Without considering a social science lens, AI risks reinforcing inequalities, perpetuating biases, and privileging dominant cultural narratives. Understanding how AI tools are adopted and adapted in diverse educational scenarios is as critical as building more efficient algorithms. No doubt, AI can capture that even experts overlook. Yet, it lacks the ethical sensitivity, cultural awareness, and critical inquiry that are essential to human learning. And perhaps that is its most important limitation. It reminds us that education is not just about processing information but about nurturing creativity, empathy, and agency. We must resist the temptation to reduce learning to what AI can calculate, without losing sight of the messy, inspiring, and deeply human aspects of knowledge-making. This SI is therefore an invitation to move beyond code and curriculum, toward the lived realities of students and the educators navigating AI-driven systems worldwide.

Our Fulbright alumni contributors examine AI in education from multiple vantage points. This SI brings together diverse perspectives from Fulbright alumni whose work spans multiple continents. Writing from the United States, Aldan Creo interrogates how AI can either bridge or deepen socio-economic divides in education. He focuses on AI bias, arguing for a more inclusive and deployment of AI tools, resisting punitive and exclusionary approaches. From the Philippines, Aliza Racelis situates the conversation in the global regulatory landscape, analyzing how different regions are adopting AI policies. Bethany Gowan reflects on how AI-enhanced learning can help students and teachers cultivate ethical literacy across cultures, equipping them to navigate complex information ecosystems. Elizabeth Radday reminds us of the hidden dangers of generative AI (GenAI), from bias to emotional dependence, and calls for ethical education starting in elementary schools. Drawing on his Fulbright experience in Thailand, Kuldeep Nagi offers a compelling perspective on

privacy and digital inclusion in contexts where face-to-face learning still dominates. Finally, Narun Pat demonstrates how cognitive neuroscience and computer science converge in AI-enabled learning, underscoring the power of global collaborations, such as those facilitated by Fulbright, in shaping equitable futures of education.

The richness of these contributions demonstrates that AI in education is not a single story, it is a tapestry woven from diverse experiences, contexts, and disciplines. By bringing together theoretical reflections, empirical studies, case-based insights, and critical commentaries, this SI highlights the urgent need for interdisciplinary dialogue. As Fulbright alumni, our contributors embody the value of crossing borders, intellectual, cultural, and geographical, in order to reimagine what education can and should be in the age of AI.

Looking ahead, the central question is not if AI will change education, but how we will guide and govern that change. The articles in this issue invite teachers, researchers, policymakers, and students to embrace critical AI literacy, to question dominant narratives, and to build systems that reflect shared values of justice, equity, and inclusion. We hope this collection sparks discussion and inspires readers to play an active role in shaping the educational landscapes of tomorrow.

"Looking ahead, the central question is not if AI will change education, but how we will guide and govern that change."

"Technology drives change; humanity defines direction."

"Beyond innovation, we need imagination."

As guest editors, we extend our gratitude to the contributors whose voices animate this SI, and to our readers, who will carry these conversations into classrooms, policy forums, and communities around the world. The future of education with AI is unwritten. Let us write it together, with courage, creativity, and conscience…

ARTICLES

Beyond The Brushstroke: Revisiting Creativity in the Age of Artificial Intelligence

Nina Maria Waals

Abstract

Emerging technologies, especially artificial intelligence (AI), transform conventional understandings of artistry and creativity. This article examines how AI-generated works challenge classical notions of authorship and property, reframing the role of human intent in the creative process. Educational and industry examples reveal the collaborative dynamics that enrich artistic endeavors, suggesting that AI can expand rather than diminish creative expression. Ultimately, broader questions about originality are raised, inviting a reassessment of art's boundaries in the digital era.

Keywords: Artificial Intelligence • Creativity • Authorship • Intellectual Property • Digital Transformation

In a quiet, dignified room of the Rijksmuseum in Amsterdam, where the air is typically thick with the reverence for classical art, an unusual exhibit prompted a subtle yet profound disquiet among its viewers. This was the unveiling of the "Next Rembrandt," a piece not crafted by human hands but calculated by the cold logic of artificial intelligence (AI). In 2016, a collaboration of technologists, historians, and artists set forth to merge the old with the futuristic, tasking themselves with creating a new work by Rembrandt, or rather, an echo of what his genius might have conceived were he alive today.

This painting, composed not from oil but from data, was algorithmically generated by analyzing the depths of Rembrandt's technique, distilled from his life's works. It presented a face that never was, yet unmistakably Rembrandtesque, challenging the observers to discern where the boundary lies between the artist's impulse and the algorithm's output. Here lay a canvas that was both a technical marvel and a cultural query, blurring the lines between man-made and machine-generated.

This moment of reveal was not just about the achievement of a technological feat; it was a deeper meditation on the essence of human creativity. It asked us whether there's a soul in the symmetry and shades managed by algorithms. In a society increasingly comfortable with the digital, the "Next Rembrandt" forces us to pause and ponder. If art is a mirror to the human condition, what

does it reflect about us when our creations are born from binary codes? This initiative is more than an artistic endeavor; it is a philosophical exploration of our age-old quest to understand what it means to create and what it means to be human in the era of AI.

Philosophical Underpinnings: Locke's Labor Theory and AI

To appreciate the contours of the current debate on AI in art, it helps to revisit the thoughts of John Locke, a philosopher who shared the temporal stage with Rembrandt. Locke's discourse on property and labor offers a classic, yet ever-pertinent framework to dissect the complexities presented by AI in the creative fields today. Locke posited in his *Second Treatise of Government* that personal ownership of property is justified through the application of labor. He argued that nature provides us with resources in common, but when an individual mixes their labor with those resources, be it tilling the land or carving a statue, those resources transform into their property.

This transformation is rooted in the expenditure of effort, which Locke viewed as a natural extension of the person themselves. Therefore, by mixing their labor with the world, individuals make a claim to ownership not just of the fruits of their labor but of a part of the world itself. The programmers and data scientists behind the "Next Rembrandt" project ventured into a realm where Locke's theory intertwines intriguingly with modern digital creations. They invested intellectual labor, analyzing Rembrandt's works, understanding his style, and translating these into algorithms that could replicate his art. If Locke were to ponder this scenario, he might consider whether these digital brushstrokes, borne from human intellect and executed by machines, constitute a new form of property. These creators did not paint, but they instructed the machine on how to do so; their intellectual efforts mixed with digital resources to create something novel and valuable. Locke's theory, thus, not only remains relevant but becomes crucial in framing our understanding of property rights in the digital age, redefining what it means to 'mix one's labor' with the materials at hand, even if those materials are as ethereal as data and code.

Distinguishing Art from Illustration: The Intent Approach

As we investigate further what can distinguish art, an integral distinction evolves between art and illustration. This nuanced perspective is insightfully broadened through the reflections of Helen Darby, a legal scholar whose expertise bridges technology, copyright law, and the arts. In a recent interview conducted with her, Darby explored how traditional concepts of authorship and moral rights are being reshaped by the advent of AI in creative fields. She eloquently underscores the significance of intent in distinguishing meaningful art from mere technical illustrations.

Picture this: a skilled artisan painstakingly replicating a renowned masterpiece, striving to mirror every brushstroke, hue, and detail. The result? Visually stunning, a perfect copy, evoking admiration from the untrained eye. But beneath this flawless facade lies a critical disparity. The forgery, despite its precision, lacks the soul of the original, the artist's intent. Each stroke, each choice, infused with the artist's vision and emotion, is absent in the imitation. Thus, while it might appear puzzling how visually identical images can be classified as art and illustration, this is not a new issue within the realm of AI-art; it is a concept as old as the art trade itself. Art, traditionally viewed through the prism of human creativity and intent, transcends mere aesthetic or functional values; it seeks to engage, provoke, and communicate. Illustration, while often skilled and functional, does not typically aspire to provoke thought or evoke deeper reflection in the same way.

This distinction becomes particularly pertinent when we consider works like the Next Rembrandt. This project, a fusion of data science and artistic legacy, challenges us to ask: Can a creation born from algorithms truly embody the intent traditionally reserved for human artists? The programmers and data scientists behind this AI endeavor did not merely feed data into a system; they engaged in a deeply creative act of teaching the machine the nuances of Rembrandt's style. Their labor was not just technical but imbued with a specific intent: to resurrect Rembrandt's artistic spirit in a new form. According to Locke, who placed great emphasis on the transformative power of labor and extending this through Darby's insights into the importance of intent, such a project does more than create a mere digital output; it crafts a new artwork, rich with human creativity and intellectual engagement.

This exploration challenges us to reconsider our definitions of art in the digital era. When AI is directed by human intent, when it is programmed to go beyond functional outputs and to imbue its creations with meaning, depth, and context, it crosses from being a mere tool to becoming a medium for artistic expression. The intent with which these machines are programmed, to create, to emulate, to innovate, becomes pivotal in distinguishing between mere illustration and genuine art. Thus, as we contemplate the evolving intersection of technology and creativity, we find ourselves engaged in a broader dialogue about the essence of human expression. The intent approach invites us to reflect not only on the outputs of our technological tools but also on the human purposes that guide them. In this way, each piece of AI-generated art serves as a mirror, reflecting back at us not just an image but a story of intent, labor, and creativity that challenges our deepest convictions about what it means to create and appreciate art in the modern world.

THE DIRECTOR'S CANVAS: AI AND THE ART OF CREATION

In the ongoing discourse surrounding the artistic validity of AI-generated works, skepticism often arises from the perception that these creations lack the personal touch, a soul behind the canvas, so to speak. However, to expand our understanding, consider the role of a film director, a figure universally acknowledged as an artist despite primarily working through the medium of instruction and collaboration. Quentin Tarantino, renowned for his distinctive cinematic style, provides a compelling parallel. Like a conductor of an orchestra, Tarantino doesn't play every instrument; rather, he directs each component to harmonize into a cohesive, artistic expression. This analogy brings us closer to appreciating the role of AI in art. When a director like Tarantino assembles a film, he sets the scene, guides the actors, selects the shots, and shapes the narrative, much as the programmers of the "Next Rembrandt" guided the algorithms to paint in a certain style. Both the director and the AI programmers engage in a high form of artistry not through direct creation but through the orchestration of numerous elements towards a singular artistic vision. Yet, one might argue that Tarantino's touch, his decisions, his style, and his vision, are palpable in every frame, something that may seem absent in a painting generated by AI.

However, this perspective may overlook the depth of human involvement in the creation of AI art. Programmers, like directors, imbue their creations with intent, style, and personal flair through the code they write and the models they build. They set the parameters within which the AI operates, much as a director controls the environment of a film set. In both cases, the art is not merely in the end product but in the vision and decision-making that lead to its creation. Recognizing this, we can see AI not just as a tool for generating art but as an extension of human creativity, reflecting the choices, style, and intentions of its human creators.

THE CREATIVE DIALOGUE: AI IN PRACTICE

The philosophical framework established by Locke's labor theory and our understanding of directorial intent finds striking manifestation in contemporary educational and industrial settings. Here, the theoretical becomes tangible, as institutions and industries engage in what might be termed a "creative dialogue" with AI, each instance reinforcing our earlier observations about the nature of artistic labor and intent.

The examination of educational environments proves particularly crucial to our discussion, as it is within these spaces that future definitions of creativity and artistic intent are being shaped. Consider the Rhode Island School of Design's integration of AI into its Master of Fine Arts program. In a thoughtful echo of Locke's theory of labor mixing with resources, students engage in a fascinating dual process: they first "labor" to teach the AI their artistic preferences and stylistic inclinations, then engage in a second

phase of creative labor as they refine and transform the AI's output. This twofold application of creative intent mirrors our earlier discussion of the "Next Rembrandt" project, but in a more intimate, individual context. The students' work becomes not merely about creating art but about developing a sophisticated understanding of their own artistic voice through its translation into algorithmic terms. Increasingly, other institutions follow similar paths. At the Berklee College of Music, for instance, students employ Google's Magenta to co-compose pieces that merge traditional music theory with algorithmic suggestions, resulting in hybrid compositions that neither the human nor the AI could have envisioned alone. This collaborative dynamic showcases how AI can facilitate unexpected creative directions, reshaping artistic practices in real time.

This educational context is particularly significant because it represents the frontier where our theoretical framework meets practical application. As these students grapple with questions of authorship, intent, and creativity in their daily practice, they are actively shaping the future discourse around AI and art. Their experiences challenge and refine our understanding of what constitutes artistic labor in the digital age, providing a living laboratory for the philosophical questions we have explored. Moreover, their work demonstrates how the integration of AI into artistic education isn't merely about learning new tools, but about developing a more nuanced understanding of creative intent itself, a key concept in our earlier discussion of what distinguishes art from illustration.

These real-world applications serve not merely as examples but as vital evidence supporting our theoretical framework. They demonstrate how the intersection of human intent and artificial intelligence creates not a diminishment of artistic labor, as some might fear, but rather a new form of creative engagement that Locke himself might have recognized as a legitimate form of property-generating labor. In each case, we see how the human element, the intent, the vision, the creative direction, remains paramount, while the AI serves as both medium and collaborator in the artistic process.

Nevertheless, these successes in education and industry raise a profound question: Are these AI-driven creations genuinely novel, or do they merely rearrange preexisting concepts in new forms? Such concerns about the boundaries of true innovation highlight the interplay between creativity and originality, how creative intent shapes the emergence of genuinely new ideas. If creativity is the capacity to generate fresh perspectives or transform existing elements in unexpected ways, then originality becomes the tangible evidence of that transformative spark. This tension naturally brings us to the next discussion, one that delves deeper into whether AI can originate truly new concepts. To address these concerns, we must explore the notion of originality.

The Illusion of Originality: AI and the Artistic Process

In her reflection on the Feist Publications vs. Rural Telephone Service case, Darby highlights a key lesson: substantial effort alone cannot guarantee copyright. The U.S. Supreme Court ruled that merely organizing phone numbers lacked the crucial element, originality, emphasizing the need for meaningful transformation. This principle resonates with current debates on AI-generated art: while AI often combines existing data, it can still create genuinely new works if guided by a transformative intent. Indeed, art's history shows no creation emerges in a vacuum; Shakespeare adapted older plots, and modern filmmakers remix past influences. Likewise, AI can remix familiar elements into novel expressions, provided human authors imbue it with creative purpose. Dismissing these results as unoriginal may overlook how art evolves through reinterpretation. Instead, we might view AI as expanding the ways we conceive of originality, forging new possibilities in the continuum of artistic transformation.

Conclusion: Reframing Creativity in the Digital Age

As we conclude this exploration of AI and artistic creativity, it becomes clear that intent remains central, both for copyright considerations and for our evolving definitions of art. Rather than focusing solely on physical skill, we now recognize a deliberate artistic vision as the core of genuine creativity. In this view, AI functions not to replace the human artist but to broaden how intent is expressed, much like brushes and chisels did in earlier eras. Still, skepticism persists. Debates on what truly qualifies as art are as old as art itself. Today, these discussions often play out on digital platforms, where some see boundary-pushing works as genuine art and others dismiss them as gimmicks. Yet the subjectivity of art means it inevitably shifts with each new medium, AI included. By embracing AI as a legitimate creative tool, we expand our idea of what is possible and push the limits of human ingenuity. Just as previous generations challenged orthodox methods, we too can reconsider and reshape creativity and authorship in this new technological landscape.

"The subjectivity of art means it inevitably shifts with each new medium, AI included. By embracing AI as a legitimate creative tool, we expand our idea of what is possible and push the limits of human ingenuity."

Further Reading

1. Bown, Oliver. Generative AI: Copyright, Ethics, and the Future of Creativity. UNSW
2. BusinessThink, February 19, 2024. https://www.businessthink.unsw.edu.au/articles/generative-AI-creative-industries-copyright-fair-use.
3. Chatterjee, A. (2022). Art in an Age of Artificial Intelligence. Frontiers in Psychology, 13.

4. Chibalashvili, A. (2021). Artificial Intelligence in Artistic Practices. Contemporary Art. 17
5. Darby, H. C. (2021). AI-created art and the moral rights of authors. Columbia Journal of Transnational Law. https://www.jtl.columbia.edu/bulletin-blog/ai-created-art-and-the-moral-rights-of-authors.
6. Santos, C. L. D., & Machado, A. R. (2020). Intellectual property on works of art made by artificial intelligence." *International Journal of Advanced Engineering Research and Science,* 7(12).

Group picture at the Naval Academy Foreign Affairs Conference (NAFAC).

BIOGRAPHY

Nina Maria Waals is a Dutch researcher specializing in the intersection of technology and global constitutionalism. She studied at Columbia University from 2023 to 2024 as a Fulbright scholar. Currently, she is a PhD candidate at the University of Amsterdam, investigating how technological advancements influence constitutional frameworks worldwide. Her interdisciplinary background and international experience inform her unique perspective on the interplay between law, innovation, and governance, aiming to deepen our understanding of emerging technologies' impact on democratic processes. She can be contacted at nninamaria.waals@fulbrightmail.org

Neuroscience Meets Artificial Intelligence in Bangkok, Thailand: Reflections from Brain Informatics 2024

Narun Pat

Abstract

As a leading program chair of Brain Informatics 2024 in Bangkok, I witnessed the power of global interdisciplinary collaboration between neuroscience and computer science. Trained as a cognitive neuroscientist with Fulbright support, I experienced firsthand how these fields converge to advance our understanding of the brain. Hosting the conference in Thailand also highlighted the country's growing potential in artificial intelligence (AI) and cognitive neuroscience, a promising sign for the region's scientific future.

Keywords: Fulbright • Thailand • Cognitive neuroscience • Interdisciplinary collaboration • Artificial intelligence

With Fulbright support, I trained as a cognitive neuroscientist at the Department of Psychology, Northwestern University (Evanston, IL) from 2009 to 2015. Cognitive neuroscience explores the neural mechanisms underlying cognitive functions. Early developments in artificial intelligence (AI) were heavily inspired by neuroscience and psychology, evident in concepts like neural networks and reinforcement learning. However, during my PhD, AI remained largely the domain of computer science, with limited integration into cognitive neuroscience. At the time, it felt only tangentially relevant to our field.

Fast forward to 2025, and AI has become more commonly incorporated into cognitive neuroscience research. For instance, our laboratory at the University of Otago in New Zealand (https://sites.google.com/view/hamneurolab/) has applied AI to predict cognitive processes from large-scale neuroimaging data. This progress has been made possible in part due to long-standing interdisciplinary collaborations between neuroscientists and computer scientists, leading to recent breakthroughs.

> ***"Our laboratory at the University of Otago in New Zealand has applied AI to predict cognitive processes from large-scale neuroimaging data. This progress has been made possible in part due to long-standing interdisciplinary collaborations between neuroscientists and computer scientists."***

Conferences: Catalysts for Neuroscience – AI Collaboration

Brain Informatics (BI) (https://braininformatics.springeropen.com/) is an academic journal that exemplifies the interdisciplinary collaboration between neuroscientists and computer scientists. Each year, Brain Informatics hosts the International Conference on Brain Informatics, which brings together experts from neuroscience and computer science fields worldwide. Since its inception in 2006, the conference has been held in various countries across the globe, including China, Canada, Japan, Poland, the UK, Italy, Australia, and the USA.

For its 17th conference, BI 2024 was held in my hometown, Bangkok, Thailand. The event was sponsored by international organizations such as the Institute of Electrical and Electronics Engineers (IEEE) Computer Society, the Web Intelligence Consortium, The International Neural Network Society, and the Chinese Society for Cognitive Science, along with private industries such as Springer and BioSemi. Local sponsors included King Mongkut's University of Technology Thonburi.

The conference featured speakers and attendees from various countries, representing both the Global South and Global North. The interdisciplinary nature of the conference attracted a wide range of expertise across computer science and cognitive neuroscience, covering topics from theoretical modeling to animal models and robots, whole-brain morphometry, graphical modeling of brain networks, and cognitive ontologies. You can view the proceedings here: https://link.springer.com/book/9789819632961.

Cultural Differences in the Conference Organization Between Computer Science and Cognitive Neuroscience

It was an honor and privilege to serve as the leading program chair of BI 2024. Although I had prior experience organizing conferences for neuroscience and psychology, BI 2024 was different. The standard practices were more aligned with computer science than neuroscience or psychology. Coordinating with committee members from Europe, North America, Asia, and Oceania, especially with the time differences, was challenging. However, it was a great learning experience.

For instance, computer scientists seem to value conference proceedings more than neuroscientists, possibly because research in computer science progresses at a faster pace. Participants, particularly those with a computer science background, are often motivated to submit a 9-12 page full paper rather than a 1,500-word abstract. These full papers were peer reviewed for originality, significance of contribution, technical merit, and presentation quality. Accepted full papers were not only presented orally at the conference but also published in a proceedings book series.

This required organizers to find reviewers who could peer review all submitted papers promptly. The reviewing process, including revisions and resubmissions, had to be completed within a couple of months from submission. This ensured that authors were notified in time to prepare for travel to the conference, which might be in a country on the other side of the world. This contrasts with traditional neuroscience/psychology journals, where there is no set timeline for article acceptance and publication. This practice also necessitated an army of experts from various fields who could quickly provide high-quality reviews. Initially, I was concerned about this, as reviewing is a volunteer work, and journal editors often struggle to find expert reviewers willing to sacrifice their weekends and nights for reviews.

In an IEEE conference like BI 2024, the group of experts who review submissions is called the Program Committee or PC. Each submission requires at least two, but ideally three, PCs to review. To my surprise, finding PCs was easier than I expected. We sent emails to previous attendees and our networks of colleagues in neuroscience/psychology and computer science. Despite their busy schedules, many accepted our invitation. They were the unsung heroes who made conferences and academic journals possible.

Given its interdisciplinary nature, another challenge was matching each submission's content to the appropriate PC's expertise. Fortunately, the conference team, including Hongzhi Kuai from Maebashi Institute of Technology in Japan, created a computerized system to facilitate this matching. As a program chair, I was grateful for their system, which allowed me to alert PCs to ensure reviews were completed on time.

Most organizers worked for the conference out of love, and we were by no means professional organizers. There were times when I had to have multiple Zoom meetings late at night to troubleshoot. This included unexpected events, such as the payment system in Thailand being flagged in some countries and last-minute personnel changes. We had to be flexible and adapt many of our initial plans. We learned a lot from this experience and were grateful for everyone's help. I believe we had a successful conference, one of the first of its kind in Thailand.

Hosting Brain Informatics 2024 Marks a Significant Milestone in Thailand's Research in Cognitive Neuroscience and AI

I completed my undergraduate degree in psychology in Thailand in 2008 before pursuing PhD training supported by a Fulbright scholarship. At the time, research in cognitive neuroscience and AI in Thailand was still in its infancy. Over the past decade, however, I have witnessed remarkable growth, driven by a new generation of Thai scientists from diverse fields including cognitive science, neurology, computer science, physics, and data science.

Many of the key contributors to this transformation, such as Sirawaj Itthipuripat, Chaipat Chunharas, Titipat Achakulvisut, Thitaporn Chaisilprungraung, Thiparat Chotibut, Itthi Chatnuntawech, and Sittiprapa Isarangura, played vital roles in organizing BI 2024. Most of them were trained abroad through Thai government scholarships, which often require recipients to return and contribute to research and education in Thailand.

Despite limited resources, this cohort has built a thriving research ecosystem, publishing in top-tier journals, forming global and local interdisciplinary collaborations, and mentoring the next generation. Their efforts, along with initiatives like BI and Braincode101 (https://braincode101.github.io/), are shaping a vibrant future for cognitive neuroscience and AI in Thailand.

Conclusion

In conclusion, I am grateful for the opportunity to be part of Brain Informatics 2024. This conference exemplifies the interdisciplinary collaboration between cognitive neuroscience and artificial intelligence. As an organizer, I learned a great deal about the cultural differences in how conferences are organized in these fields. Having witnessed firsthand the promise future of cognitive neuroscience and AI, I am hopeful for the continued advancement of research in these areas, both in Thailand and globally.

Further Reading

1. Buianova, I., Silvestrin, M., Deng, J., & Pat, N. (in press). *Multimodal MRI Marker of Cognition Explains the Association Between Cognition and Mental Health in UK Biobank*. eLife. https://doi.org/10.7554/eLife.108109.1
2. Itthipuripat, S., Ascoli, G. A., Li, A., Pat, N., & Kuai, H. (Eds.). (2025). *Brain Informatics: 17th International Conference*, BI 2024, Bangkok, Thailand, December 13–15, 2024, Proceedings, Part I (Vol. 15541). Springer Nature. https://doi.org/10.1007/978-981-96-3294-7
3. Pat, N., Wang, Y., Bartonicek, A., Candia, J., & Stringaris, A. (2023). Explainable machine learning approach to predict and explain the relationship between task-based fMRI and individual differences in cognition. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhac235
4. Tetereva, A., Knodt, A., Melzer, T., van der Vliet, W., Gibson, B., Hariri, A., Whitman, E., Li, J., Khakpoor, F. L., Deng, J., Ireland, D., Ramrakha, S., & Pat, N. (2025). Improving Predictability, Test-Retest Reliability and Generalisability of Brain-Wide Associations for Cognitive Abilities via Multimodal Stacking. PNAS Nexus. https://doi.org/10.1093/pnasnexus/pgaf175

5. Wang, Y., Anney, R., & Pat, N. (2025). *The relationship between cognitive abilities and mental health as represented by cognitive abilities at the neural and genetic levels of analysis*. eLife. https://doi.org/10.7554/eLife.105537.1

During the opening ceremony of Brain Informatics 2024, I was honored to receive a gift from Dr. Krittika Tanprasert, in recognition of my contributions as a leading program chair. Dr. Tanprasert, Director of the Learning Institute at King Mongkut's University of Technology Thonburi, represented the host institution of this conference.

BIOGRAPHY

Narun Pat, PhD, is a senior lecturer at Department of Psychology and Neuroscience Teaching Program at University of Otago—Dunedin, New Zealand and a former Fulbrighter to the US at Northwestern University, Evanston, IL. He applied AI to predict cognitive functioning from large-scale brain-scan data. He can be reached at narun.pat@otago.ac.nz.

The Ethical Issues Surrounding Artificial Intelligence-Powered Learning

Kuldeep Nagi

Abstract

Over the past three years, Artificial Intelligence (AI) has rapidly transformed education, reshaping how students learn and how institutions deliver services. Ethical AI use requires enhancing learning while protecting dignity, privacy, and equity. Educators and developers must ensure AI tools promote inclusive, student-centered environments without reinforcing inequalities or compromising rights. This article also draws on the authors' Fulbright experiences in Thailand, where traditional face-to-face teaching still predominates, offering insights into AI's ethical integration.

Keywords: Artificial Intelligence • Education • eLearning • Ethics.

Since its inception in 2019, more than a dozen major companies have been developing Artificial Intelligence (AI)-based innovations implemented as multiple products, services, apps, and bots. More recently a Chinese startup known as Hangzhou DeepSeek Artificial Intelligence (HDAI) Basic Technology Research Co., Ltd introduced a new variant of AI called DeepSeek which has rattled companies in the USA and elsewhere. Russia, UAE (Dubai), India, the UK, and other members of the EU are also developing their own AI models. Although companies are pouring billions of dollars into AI research, many myths and speculations about the dangers of AI have dominated the media. The abuse of AI by students has also made appalling headlines.

Many people in academia believe that AI will replace teachers, reduce creativity, or result in a one-size-fits-all approach to learning. The reality is that AI is a tool designed to enhance, not replace, the role of educators, offering personalized support and freeing up teachers to focus on more meaningful interactions with their students. Similarly, concerns about AI stifling creativity overlook its potential to inspire innovative teaching strategies and foster student engagement.

Another common myth in academia is that AI will promote plagiarism and hamper creativity. Similar claims have also been made about search engines, Turnitin, Grammarly and social media platforms. Another popular myth is that AI will hamper skill development. On the contrary, AI usage depends on the knowledge and skills of the users.

AI is rapidly transforming education, yet its adoption is often accompanied by misconceptions that hinder its full potential. Just like the internet and social media, AI is also prone to unethical use and abuse. Hence, training of its users is a key concern. Although there is always a possibility of abuse, it is important to consider AI's ethical implications in education, such as data privacy, bias, discrimination, autonomy, and surveillance.

Data, Privacy and Power — Evolving Regulations for AI

The development of ethical principles, policies, and regulations for AI is a crucial area of focus in today's rapidly advancing technological landscape. Key issues like surveillance, privacy, autonomy, bias, discrimination are being addressed by several civic organizations, especially in law enforcement. For example, the development of the social credit system in China and the ban on facial recognition technology in San Francisco illustrate the divergent approaches taken in different regions toward the ethical use of AI and the balance between economic growth and personal privacy concerns.

International coordination for creating standards is necessary to address AI's unique ethical, legal, and philosophical challenges. Several international organizations, such as UNESCO, OECD, the IEEE, and the EU have developed AI ethical frameworks. These frameworks emphasize principles of human rights, accountability, transparency, and fairness.

Since 2006, traditional publishers, such as Pearson's MyLab, have combined in-class and online learning, providing simulations, videos, quizzes, and other resources to enrich the student experience. As of June 7, 2025, more than 125,000 studies listed by Google Scholar indicate AI's positive effects in education. A quick review of these papers indicates that regulation of ethical use of AI in education is rapidly evolving.

Contextualizing AI in Thai Education

As an eLearning expert and Fulbright Fellow (2005), I have closely followed shifts in Thai education since the onset of COVID-19 in late 2019. The pandemic forced rapid adoption of online learning platforms like Moodle and Microsoft Teams, along with widespread use of Zoom, LINE, Skype, WhatsApp, and other social media tools. After the pandemic subsided in 2023, the system has largely reverted to traditional face-to-face (F2F) teaching. A recent study highlights critical challenges: outdated rote curricula, limited teacher training, low digital literacy, and poor STEM and English performance. These issues contribute to weak results in global assessments like PISA and TIMSS. An aging population has also led to declining student enrollment. Although the Ministry of Education's 2024 policy (No. 2-10) calls for integrating AI and enhancing digital and language skills, implementation remains limited. Despite significant education funding, questions persist about

its effectiveness. For AI to succeed in Thai education, the country must first tackle digital literacy gaps, ensure ethical use, and promote inclusive access. With the right policies, AI has the potential to transform learning for both students and educators.

Conclusion – AI is here to stay

With the rapid rise of AI-powered solutions, there is growing concern that both teachers and students may become overly dependent on this new technology. Over time, such reliance could lead to significant mental and social challenges. Similar concerns about search engines and online learning have been raised, though such skepticism has sometimes hindered meaningful reform in the Thai education system.

It is becoming increasingly evident that AI will take over many of the routine tasks currently handled by educators. However, this does not diminish their role. On the contrary, it enhances their value. Students still need something AI cannot offer- genuine human connection.

Today, AI excels at managing personalized learning and skill development, and at providing adaptive diagnostics for tasks that once consumed a great deal of educators' time and energy. This shift allows instructors to evolve into "Guides" who do not merely deliver content, they lead, observe, and respond. They help students discover their potential, something AI cannot replicate, and likely never will. Therefore, educators must be trained not just to use AI, but to teach students how to engage with it purposefully, not as a shortcut, but as a tool. However, if we fail to leverage it during creative, analytical, or innovative tasks, we are missing valuable opportunities.

AI as co-pilot, the instructor as guide, and the student as leader, this is the future of education. AI will not replace teachers; it will replace the aspects of teaching that never truly required a human touch. What keeps great educators essential is not their ability to deliver content or grade assignments, it is their irreplaceable humanity. In today's world, that is exactly what students need most. Even Microsoft founder Bill Gates is optimistic about AI's potential to reshape learning. *"In the next five to ten years, AI-powered software will transform how we teach and learn,"* he predicts.

> ***"AI as co-pilot, the instructor as guide, and the student as leader – this is the future of education. AI will not replace teachers; it will replace the aspects of teaching that never truly required a human touch."***

FURTHER READING

1. Nagi, Kuldeep, New Social Media and Impact of Fake News on Society (June 6, 2018). ICSSM Proceedings, July 2018, Chiang Mai, Thailand, pp. 77-96, Available at SSRN: https://ssrn.com/abstract=3258350
2. Numor, P. (2024). Thai government gazette (on 13 November, 2024), notifications of Ministry of Education Re : Education Policy of the fiscal year 2025-2026. *Ministry of Education Notification.*
3. Noah, S. (2025). Bill Gates Paints a grim picture for future of work. Retrieved from https://jasondeegan.com/only-three-jobs-will-survive-ai-bill-gates-paints-a-grim-picture-for-the-future-of-work/
4. Sarasean, T. (2024). Challenges and Opportunities: Reforming the Thai Education System for Global Competitiveness. *Journal of Exploration in Interdisciplinary Methodologies (JEIM)*, 1. Retrieved from https://so19.tci-thaijo.org/index.php/JEIM/article/view/741
5. Thabmali, P., & Chakamanont, S. (2025). Guidelines for enhancing Thai Teacher's AI literacy in the digital age. *Journal of Education and Innovation*, 27(1), 27-50.

My Book was given to the Director of the library at Assumption University for sharing with the library patron.

BIOGRAPHY

The author is retired faculty of Seattle Community College District (SCCD), WA, USA. He received the Fulbright Scholar Award in 2005 and the Dan Evans Award for Excellence in Teaching for his pioneering work in eLearning. He has written dozens of academic and research papers for IEEE-sponsored conferences and many highly indexed international journals. Some of them are listed in Academia.com, Google Scholar, Scopus, and Social Science Research Network (SSRN), where he is also listed among the top 10 authors. He has written two books about the history and evolution of eLearning. He specializes in eLearning, EdTech, Social Media, and Higher Education research.

ARTIFICIAL INTELLIGENCE AND EDUCATIONAL INNOVATION IN AFRICA: OPPORTUNITIES, CHALLENGES, AND ETHICAL CONSIDERATIONS

HOUSSINE SOUSSI

ABSTRACT

AI is reshaping education globally, and Africa is actively part of this transformation. This article examines innovative uses of AI in African education while addressing challenges of equity, access, and ethics. Drawing on my experience as a Fulbright Scholar from Morocco with the African Studies Initiative at the University of Minnesota, I argue that AI's success in Africa hinges on strong policy frameworks, digital infrastructure, and ethical considerations that prioritize inclusivity and fairness.

Keywords: Artificial Intelligence • Education • Innovation • Africa.

I was honored to be asked to speak about artificial intelligence (AI) and innovative teaching methods in Africa by the University of Minnesota's African Studies Initiative (ASI) as a Fulbright Scholar from Morocco. In addition to examining the connections between Africa and the US, particularly in the areas of education and knowledge production, the ASI seeks to elevate African perspectives in international discourse. The purpose of my talk was to provide a thorough analysis of the application of AI in African schools, covering both its advantages and ethical concerns.

Working with ASI taught me many valuable lessons, one of which was the significance of contextualizing AI applications within Africa's diverse educational landscape. AI has the potential to reduce learning gaps and facilitate access to high-quality education, but its implementation must take into consideration infrastructure issues, disparities in income and educational attainment, and cultural relevance. During discussions with faculty and students, I explored how AI-powered solutions in Africa differ from those in developed countries due to the continent's unique technological and economic realities. African-led AI innovations that prioritize inclusivity and sustainability are essential, as demonstrated by our research's focus on global-local connections in knowledge production.

Reframing The Narrative – Africa Beyond Challenges

For far too long, conversations about Africa have been dominated by problems like poverty, underdevelopment, and conflict. These facts are unavoidable, but they are only one part of the whole picture. Africa is also a continent full of young people who have untapped potential and are becoming more open to trying new things. We need to change the narrative if we want to understand how AI can change things.

Africa is the continent with the youngest population, with over 60% of its citizens under 25. This sizable youth population presents an excellent opportunity to improve education through the use of technology. Unlike older populations in other parts of the world, young people in Africa are highly adaptable and eager to use new technologies. AI can assist young people in obtaining the information they require to improve their futures, in conjunction with the proliferation of mobile devices and digital innovation.

Mobile technology is already very common in Africa. The GSM Association says that by 2025, there will be 615 million unique mobile subscribers in Sub-Saharan Africa. This technology makes it possible for AI-powered learning platforms to reach even the most remote communities. Africa is ready to not only use AI, but also to use it in ways that are unique to its needs.

It is especially exciting that Africa's AI ecosystem is already starting to take shape. Data Science Nigeria and Data Science Africa, as well as Makerere University in Uganda, are helping to develop local talent. Researchers like Rose Nakasi are making new AI apps that solve problems that are unique to Africa, like the detection of malaria. These projects are more than just new technologies; they are part of a growing movement that is changing how Africa uses AI.

The Continental AI Strategy of the African Union, which was approved in July 2024, sets out a vision for an Africa-centered, development-focused approach to AI. This plan stresses how important it is to make AI work for everyone, be fair, and sustainable. It imagines an Africa where AI helps everyone, from urban centers to rural villages, from the privileged to the marginalized. This is the future we need to work toward.

AI as a Catalyst For Educational Transformation

Education is still one of Africa's biggest problems. Millions of students don't have easy access to quality learning resources and materials. AI offers a unique chance to close this gap by offering technology-based solutions that can be used anywhere and by anyone, regardless of location or income level. Kenya, Rwanda, and Ghana are already trying out AI-powered educational tools that make learning more personal, support educators, and improve digital literacy.

One of the most important ways that AI is changing education in Africa is by making learning more personalized. *Eneza Education* in Kenya is changing the way people learn from a distance by using AI-powered chatbots to send lessons that are in line with the curriculum to simple mobile phones. Studies have shown that students who actively use the platform show an average improvement of over 22% in their scores. The platform has reached over 10 million learners in several African countries. This method makes sure that even students in areas with poor internet access can get a good education. *Ubongo*, an edutainment company based in Tanzania, also uses AI-driven adaptive learning models to make educational content more suitable for young learners all over Africa. Its animated shows have been seen by more than 25 million homes across the continent. Independent evaluations have shown that its young viewers have made big improvements in their reading, writing, and social-emotional skills.

AI has also been very important in helping students with disabilities. In Ghana, *LearnerX* and other platforms use AI-powered speech recognition and text-to-speech apps to help students who have trouble seeing or hearing. These tools not only make things easier to get to, but they also make sure that learning spaces are more welcoming to everyone. These AI applications are in line with constructivist learning theory because they encourage learning that is active, personalized, and relevant to the situation. Vygotsky's Zone of Proximal Development (ZPD) is particularly relevant, as AI tools support student learning through real-time evaluation of skills and advancement.

AI-Assisted Teaching And Professional Development

Beyond benefiting students, AI also plays a crucial role in supporting teachers. A lot of African countries face severe teacher shortages, leading to high student-teacher ratios and limited individualized instruction. AI-powered tools are being used to improve teacher training, take care of administrative tasks, and give teachers feedback on their teaching in real time. AI takes care of administrative tasks like grading and keeping records, which lets teachers focus on higher-order teaching practices like helping students think critically and giving them social and emotional support, as frameworks like Bloom's Taxonomy emphasize.

As part of its *National AI Policy*, the Rwandan government has embraced AI-driven teacher development initiatives. Teachers can improve their methods by using AI analytics platforms that track classroom interactions and produce reports on how well they are teaching. AI can increase pedagogical effectiveness while lowering teacher workload, as demonstrated by initiatives like *SmartClass*, which gives teachers AI-generated suggestions on how to deliver lessons. Additionally, AI-powered tools evaluate classroom interactions and give teachers immediate feedback on their methods and students' participation. For instance, virtual teaching assistants and AI chatbots are assisting teachers with administrative and grading tasks so they can concentrate

more on mentoring and student engagement. Teachers can improve their teaching strategies and increase the effectiveness and interactivity of learning by utilizing AI. I noticed conflicting opinions among educators during my fieldwork and Fulbright exchange. Some worry about the lack of training and fear of being replaced, while many are excited about AI's potential to improve student outcomes and reduce workload. For example, to increase confidence in the adoption of AI, faculty members at the *University of Ibnou Zohr* in Morocco stressed the necessity of focused professional development.

CASE STUDIES: AI IN EDUCATION ACROSS AFRICA

A number of African countries have made significant progress in using AI in their schools. Looking at case studies from Kenya, Rwanda, Ghana, and Morocco can help us understand the successes and challenges of using AI to improve education.

KENYA: AI-POWERED MOBILE LEARNING

Kenya has become a leader in AI-powered education, especially through mobile learning solutions. *Eneza Education* is an AI-based ed-tech platform that sends students in rural areas SMS-based lessons that are in line with their schoolwork. Eneza's AI-driven platform has more than six million users in Africa. It customizes content to each student's learning pace and includes quizzes, personalized feedback, and progress tracking.

M-Shule is another example. It is Kenya's first AI-powered SMS learning platform that sends personalized educational content to primary school students. The platform looks at how students respond and changes lesson plans to fit each student's needs, making sure that each student has their own learning path. These solutions have been very helpful in solving Kenya's education problems, especially in areas where there aren't many qualified teachers.

RWANDA: GOVERNMENT-LED AI POLICY AND TEACHER TRAINING

By incorporating AI into its national education policies, Rwanda has set itself apart. Introduced in 2023, the nation's *National AI Policy* highlights the use of AI in curriculum development, digital learning, and teacher preparation. For example, the *SmartClass* initiative uses analytics driven by AI to evaluate teacher performance and provide customized training based on classroom interactions. In order to ensure that teachers are prepared to incorporate AI tools into their teaching strategies, the government also works with international organizations like UNESCO to implement AI-driven teacher education programs.

Furthermore, organizations like the *African Institute for Mathematical Sciences* (AIMS), which uses AI-driven research tools to improve STEM education, are prime examples of Rwanda's use of AI in higher education. Rwanda has established itself as a leader in the adoption of AI in education thanks to its emphasis on curriculum improvement and AI-driven teacher training.

Ghana: AI for Special Education and Skills Development

Ghana has made a lot of progress in using AI in vocational training and special education. The *LearnerX* program, which was started with the help of Ghana's Ministry of Education, uses AI to help students with disabilities learn. The platform uses speech recognition and natural language processing to provide real-time transcription services and personalized reading help, making it easier for students with vision and hearing problems to learn.

AI is also being added to Ghana's vocational training programs. The *AI-powered Skills Development Hub*, which was built with IBM's help, offers personalized learning experiences that are based on students' career goals. The platform looks at trends in the job market to suggest relevant courses and offers AI-based mentorship to make sure that students learn the skills that are in demand in the job market.

Morocco: A Growing Hub For AI And Education

Morocco is becoming a major force in AI-powered education. In order to promote research and innovation in AI applications for education, the government has started academic programs at universities that are specifically focused on AI. Leading the way in AI research, *Mohammed VI Polytechnic University* (UM6P) has been creating tools to improve AI-powered tutoring programs and online education.

Challenges And Ethical Considerations

While AI presents a powerful tool for educational transformation in Africa, its implementation comes with significant challenges that must be addressed to ensure equity, fairness, and responsible deployment.

Bridging the Digital Divide

The digital divide, which continues to be a major obstacle for AI-based education, is among the most significant issues. While mobile phone usage has increased, disparities in digital literacy, internet access, and technology infrastructure persist, particularly in underserved and rural areas. AI-enhanced education has the potential to worsen rather than improve educational inequality if not all people have equal access to these resources. According to the Continental AI Strategy of the African Union, we must make targeted

investments in affordable technology, infrastructure, and national policies that promote digital inclusion if we are to close the digital divide. Initiatives like Smart *Africa* seek to increase access to digital resources, but more sustained investment and collaboration between governments, corporations, and international organizations are required to make AI accessible to all.

Data Privacy and Security

Data privacy and security are also very important issues. AI systems need a lot of student data to make learning more personal and improve student performance. But this dependence brings up important moral issues: How do you get and use student data? Are privacy rights being protected well enough? It is very important to have strong data protection frameworks, especially when working with vulnerable groups like children and people who are on the outside of society. The African Union's Data Policy Framework calls for stronger policies for governing data. It asks countries to put in place protections for personal information while also encouraging openness and trust in AI systems. African countries need to establish strict rules about how data can be used, stored, and shared to ensure that AI is used ethically. They also need to teach students, parents, and educators about their rights in the digital age.

Bias and Algorithmic Fairness

AI systems can only be fair and unbiased if the data they are trained on is fair. If datasets show that there are already differences in education and society, AI tools could unintentionally reinforce those biases, which would be unfair to students. For instance, AI-powered learning platforms may focus on content that fits with urban school systems and ignore knowledge systems from rural or indigenous areas. From the student's point of view, this is not just a theoretical problem; it is a profoundly demotivating experience. If educational tools don't match a student's language, culture, or real-life experiences, they may feel left out of the learning process and more marginalized. To make sure that algorithms are fair, we need to make datasets that are inclusive and representative of Africa's linguistic, cultural, and educational diversity. The African Union's AI Strategy says that AI models should be made and tested in ways that get rid of systemic biases and promote fairness.

When I talked to students at the National School of Business and Management of Dakhla (in southern Morocco), it was clear that they are most interested in learning when educational technologies are culturally and linguistically appropriate. Several students said they were frustrated with AI tools that automatically show content in French or English, which doesn't always match their native languages or the places they live. Some people said that examples and case studies often talk about situations in other countries or cities, which makes the material feel like it's not related to their own lives.

These thoughts show how important it is to make AI systems that are not only technically sound but also useful and accessible to everyone in the area. For AI applications to be fair and open to everyone, it is important to develop AI locally, where African researchers and engineers with a deep understanding of the region build solutions.

Teacher Training and AI Adoption

To use AI-powered tools effectively in the classroom, teachers need to be trained in how to do so. Many teachers in Africa have not worked with AI-driven technologies before, and they might have trouble using these tools in their classrooms if they don't receive enough training. Governments and schools need to spend money on ongoing professional development programs that teach teachers how to use AI tools. This will help them use AI to improve their teaching instead of replacing it. For instance, Rwanda has included AI training in its programs for teacher education so that teachers are ready to use AI in the classroom. For AI to be successfully integrated into African education systems, it is important to scale up these kinds of projects across the continent.

Localization and Cultural Context

African languages, educational systems, and cultural contexts must all be taken into consideration when developing AI solutions. Many AI-driven educational tools are developed outside the continent, often without consideration for local realities. AI-powered learning platforms should make it easier for students to learn in more than one language, especially in places where students speak more than one language at home and at school. For instance, Morocco is using AI-powered language learning platforms to meet the needs of its diverse language environment, which includes Arabic, Berber, and French. By integrating AI tools with national curricula and pedagogical practices, the development of context-aware AI systems guarantees that learning stays applicable and accessible to African students.

Addressing AI Mistrust and Mindset Adaptation

In addition to infrastructure and policies, AI adoption in education necessitates a transformation in the mindset of educators, governments, and communities. People may not want to use AI widely because they do not trust it and are worried about losing their jobs and becoming too reliant on technology. Many people are worried that AI will take the place of human teachers, but AI is actually a tool that helps teachers do their jobs better. To get people to have a positive view of AI, there need to be public

awareness campaigns, open discussions, and policymaking processes that include teachers, students, and parents. By clearing up misunderstandings and showing how AI can be useful in real life, stakeholders can gain trust and faith in AI's role in education.

Policy Implementation and Regulatory Frameworks

In order to get the most out of AI in education, we need strong policy frameworks. Rwanda is the first African country to have a national AI policy that focuses on education. This sets an example for other countries to follow. AI regulation needs to find a balance between innovation and oversight to make sure that AI technologies are developed in a way that is fair to everyone and that they stay accessible. Governments should set clear rules for AI in education that make sure it is used correctly and encourage responsible development. Through regional AI governance initiatives, African countries can work together to make sure that policies are in line with each other. This will create a strong AI regulatory ecosystem that encourages responsible AI use and new ideas.

Conclusion

> ***"My time on the Fulbright program has taught me a lot about how AI can change education for the better. Working with AI-based learning models in the U.S. has helped me think critically about how they might work in Africa."***

My time on the Fulbright program has taught me a lot about how AI can change education for the better. Working with AI-based learning models in the U.S. has helped me think critically about how they might work in Africa. The significance of localized AI solutions, tools created with consideration for African languages, cultural contexts, and socioeconomic realities, is among the main lessons learned. To ensure relevance and efficacy, it will be essential to develop AI education platforms internally rather than depending on technologies that are developed outside the country.

Also, it is important for academic institutions, policymakers, and the private sector to work together. AI's role in education should not just be to implement technology. It should also include teacher training programs, digital literacy initiatives, and ethical guidelines to get the most out of it while lowering the risks.

Further Reading

1. Chisom, O. N., Unachukwu, C. C., & Osawaru, B. (2023). Review of AI in education: Transforming learning environments in Africa. International Journal of Applied Research in Social Sciences, 5(10), 637–654. https://doi.org/10.51594/ijarss.v5i10.725

2. Ezumah, B. A. (2020). Critical Perspectives of Educational Technology in Africa, Design, Implementation, and Evaluation. Springer. https://link.springer.com/book/10.1007/978-3-030-53728-9
3. Maisiri, J., & Musonza, S. (2025, July 20). *The cultural cost of AI in Africa's education systems*. UNESCO. https://www.unesco.org/en/articles/cultural-cost-ai-africas-education-systems

Embracing the Fulbright spirit of connection at the Fall 2024 Cleveland Fulbright Scholar Seminar.

BIOGRAPHY

Dr. Houssine Soussi is an Associate Professor of Business Communication and Head of the Management Department at the National School of Business and Management of Dakhla, University of Ibnou Zohr, Morocco. As Chair of the International Conference on Entrepreneurship, Creativity, and Innovation Management (ICECIM) and a member of the executive committee of the Africa Business and Entrepreneurship Research Society (ABERS), Dr. Soussi has hosted and led prestigious conferences that bring together thought leaders and experts from across Africa and beyond. He received a 2024-2025 Fulbright Scholar award. He can be contacted at h.soussi@uiz.ac.ma and ORCID: https://orcid.org/0000-0002-4708-952X

Ethical Considerations in the Adoption of AI for Education: A Social Science Perspective

Godwin Chukwuka Enwu and JohnBosco Chika Chukwuorji

Abstract

Artificial Intelligence (AI) is transforming educational and research environments through innovative tools and methodologies. However, these advances are accompanied by ethical concerns regarding bias, privacy, and fairness. This article explores the implications of AI integration in social science education and research, identifies emerging ethical challenges, and proposes strategies for responsible and equitable AI use from the Fulbright lens.

Keywords: Artificial Intelligence • Digital divide • Education • Ethics • Innovation.

Artificial Intelligence (AI) is described as the ability of a digital computer or computer-controlled robot to perform tasks that are ordinarily associated with intelligent beings. The term is commonly applied to systems endowed with the intellectual processes characteristic of humans, such as meaning, generalization, or learning from experience. Since the inception of AI in the 1940s, digitalized computers have been programmed to carry out many convoluted tasks such as discovering proofs for statistical and mathematical theorems. From the early focus on simulating human intelligence and expert systems, AI has moved toward data-driven approaches that leverage machine learning, neural networking, and deep learning. The use of AI in social science educational research is embedded in two main strands: (1) the development of AI-based tools for classroom teaching and (2) the use of AI to measure and assess growth in learning. Some of the examples of AI learning include, GPT-4 and adaptive learning platforms like Khanmigo, which support both educational practices and policy analysis.

AI holds important promise for social science education as it offers innovative tools and approaches to improve teaching and learning. Educators and researchers that leverage the advantages offered by AI can personalize instructions, provide real-time feedback and develop adaptive learning ecosystem. When AI-driven platforms personalize learning experiences, they effectively address the diverse needs of students and reduce inequalities. AI tools can analyze social media posts, enabling policy makers to track and digest public sentiments regarding topical issues like climate change or public health interventions. For example, researchers can make use of machine learning to predict areas at high risk of homelessness, guiding resource allocations and intervention strategies. In disaster response, AI algorithms may be used

to analyze data from various sources, thereby optimizing responses during natural disaster, enhancing response time and reducing harm and destruction. AI chatbots provides mental health support, which allows users to access assistance at any time while collecting data to improve services.

Ethical Issues Related to AI in Social Science Education and Research

AI offers exciting possibilities in education, but its applications come with obvious limitations. AI in education and research presents numerous ethical issues, as these technologies can have great influence on both processes and outcomes. AI system can perpetuate biases, which ultimately leads to unfair outcomes. For instance, there are apprehensions that the use of AI in marking assessments may be discriminatory towards students from less privileged or marginalized backgrounds. There is an ongoing digital divide that affects student access to technological devices, which varies significantly from one region to another and from different socialeconomic backgrounds. This raises critical concerns regarding the fairness and inclusiveness of educational outcomes driven by AI systems.

The integration of AI technologies in social science educational research had broadened the existing inequalities in data collection, as institutions with a large financial base can leverage their financial strength to secure more data than their counterparts with less financial capacity. Underfunded institutions struggle to measure up, which further exacerbates disparities in research quality and access. AI models often seek to find patterns across large datasets which ultimately leads to overgeneralization. This puts at risk oversimplifying of complex social dynamics and neglecting unique local factors that are very relevant to understanding specific communities or certain populations.

AI always relies on massive datasets, which includes personal data to drive insights. This raises questions about individual privacy and the potential for abuse of sensitive information. The ethical burden lies in balancing the need for data collection against individual rights for privacy. AI systems have the capability to operate as something whose internal workings are unknown or hidden, while its inputs and outputs are known. This makes it difficult to understand how decisions or postulations are made, and lack of transparency can hinder accountability for research outcomes.

Obtaining informed consent can result in a cumbersome process when AI systems analyze data not initially intended for research purposes. This impedes on informed consent. AI applications in educational research also manipulate or impede the autonomy of participants, especially if they are not fully aware of their engagement in research process.

AI tends to focus on quantifiable aspects of social phenomena, potentially overlooking qualitative dimensions that are crucial in social science research. Reductive methodologies might ignore the complexities of human behaviors, cultural contexts, and social nuances which ultimately lead to incomplete or misleading interpretations. An overreliance on AI can diminish critical thinking skills among researchers and may lead to devaluation of traditional research methods. Scholars may become dependent on AI assistant tools for data analysis which could stifle innovations in qualitative research and human-based solutions. There is also long-term societal impact of AI. The development of AI in social science research offers the promise of efficiency but might lead to societal changes that may affect employment, societal dynamics, and governance structures.

"The development of AI in social science research offers the promise of efficiency but might lead to societal changes that may affect employment, societal dynamics, and governance structures."

STRATEGIES TO MITIGATE ETHICAL CONCERNS IN THE APPLICATIONS OF AI IN SOCIAL SCIENCE EDUCATION AND RESEARCH

There are several ways to mitigate some of the ethical issues that plague the application of AI in social science research.

Explainable AI: There should be detailed explanation of AI systems in understandable and transparent ways. Implementing explainable AI techniques and AI decisions in education and research allows social scientists to articulate the rationale behind AI decisions.

Data Privacy: Privacy protection helps to enforce strong data governance and compliance. For instance, federated learning enables AI model to learn from data without transforming sensitive data or pieces of information while protecting participant's privacy. Data protection and data security practices such as data encryption will ensure that sensitive information collected in social science research is properly and carefully stored and transmitted. This helps in protecting participant's identities and personal information against attacks and unauthorized access.

There are several ways to mitigate some of the ethical issues that plague the application of AI in social science research.

Explainable AI: There should be detailed explanation of AI systems in understandable and transparent ways. Implementing explainable AI techniques and AI decisions in education and research allows social scientists to articulate the rationale behind AI decisions.

Data Privacy: Privacy protection helps to enforce strong data governance and compliance. For instance, federated learning enables AI model to learn from data without transforming sensitive data or pieces of information while protecting participant's privacy. Data protection and data security practices

such as data encryption will ensure that sensitive information collected in social science research is properly and carefully stored and transmitted. This helps in protecting participant's identities and personal information against attacks and unauthorized access.

Inclusivity: Fairness in the development and use of AI promotes algorithms that reduce bias. For instance, AI based recruitment tools can be adjusted to ensure that a particular population demographic is not favored more than the other group. This helps to achieve fairness during hiring practices. Inclusivity in the development and use of AI helps to engage diverse stakeholders in AI development. This practice obtains in healthcare, involving diverse communities while developing AI system for forecasting health outcomes to engender responsiveness to resolving different cultural needs. Collaborative development of AI also helps to foster collaboration within and across disciplines. A partnership between data scientists and social scientists can improve the growth and development of ethical systems ensuring that diverse community needs are being achieved.

Cybersecurity: Cybersecurity can be of immense help in guarding against ethical anomalies that is associated with AI in education and research using AI smart assistance tools. Cybersecurity can help in auditing and monitoring of AI algorithms for bias. This includes analyzing the datasets used for training AI algorithms to ensure that they are diverse and representative, thus minimizing ethical anomalies in research outcomes. Through regular updates, AI systems can be optimized with the latest security patches. This helps mitigate any biases that may pop up from outdated models or technologies. This ensures that they are aligned with current ethical principles. Security of all AI systems ensures that AI systems are secured against cybersecurity attacks and are prevented from malicious altercations that can lead to unethical research outcomes. Adhering to cybersecurity regulations and standards ensures that ethical principles are followed in data collection and applicability. Embedding cybersecurity practices within methodologies in education and social research helps to uphold ethical standards while at the same time harnessing the potentials of AI technologies.

Accountability: Traceability practices can ensure that all actions taken by AI systems are traceable, and this is vital for accountability. This allows researchers to identify and correct unethical behaviors or biases within AI systems. Compliance helps to guard against the breach of participants rights and helps to increase research integrity.

By adopting these strategies, the potential ethical issues in the use of AI would be greatly minimized, and tasks will be performed with a high level of precision and at a reduced cost.

CONCLUSION

I (JohnBosco Chika Chukwuorji) received funding from Fulbright to conduct my research in the US because the technology and expertise for my proposed psychophysiological research on emotion regulation and depression were not accessible in Nigeria at that time. At Cleveland State University (CSU), I rigorously trained for my research ethics certification. I was also exposed to digital platforms for public presentation, psychometric evaluation, online survey administration, and statistical analysis. This broadened my methodological toolkit in behavioral health research. Immersed in an academic environment that emphasized global collaboration and cutting-edge methods, I was able to publish my papers, as reflected in the 100% increase in my research outputs in 2021, which was sustained, in 2022. Exposure to these emerging technologies reinforced my value for open science, digital ethics, and the use of technology in teaching, research and clinical work (e.g., mobile health tools or AI-assisted diagnostics). Upon my return to Nigeria after the Fulbright program, the gaps in my access to technology were filled by my continued collaborations with researchers in the Western industrialized countries. In sum, my Fulbright journey reinforced my conviction that technological innovation, including AI, must be grounded in ethical responsibility, human dignity, and social justice.

As the Fulbright program often exposes grantees to diverse educational environments, international collaboration, and interdisciplinary perspectives, we have a few recommendations for Fulbright alumni working at the AI-education interface. First, they should champion equity-centered AI in education by leveraging their global connections to advocate for development of, or lend their voice in promoting, AI tools that are inclusive and adaptable to low-resource settings, multilingual classrooms, and culturally diverse learners. Fulbright alumni should partner with local educators and communities to ensure AI applications reflect contextual needs and values. I believe that AI systems will become more meaningful and impactful when its access and use not only becomes more equitable, but also culturally sensitive and contextually informed.

Having recognized that AI could be a double-edged sword, alumni should support or create open-access educational platforms powered by AI that address barriers like language, disability, or connectivity. Second, they should use their networks to strengthen cross-cultural research on the use of AI in teaching and research by conducting comparative studies on the impact of AI tools (e.g., adaptive learning platforms) across educational systems in the Global South and Global North. I worked in two research laboratories at CSU and my contact with other students from different countries broadened my appreciation for global standards in education, student-centered learning, and the importance of cross-cultural understanding in research and teaching.

Third, as educators, policy makers or implementers, it is their responsibility to promote ethical literacy around AI by organizing workshops or webinars for students and educators on digital ethics, algorithmic bias, data privacy, and responsible use of AI. In their courses, alumni who are lecturers should consider embedding AI ethics modules in relevant coursework. Fourth, as many institutions are gradually developing guidelines around the use AI, Fulbright alumni should get involved in the important work of ethical procurement and implementation of AI tools in schools and universities. Their emphasis should be on promoting frameworks that protect learner autonomy, ensure fairness, and uphold human rights.

AI has become an emerging powerful tool that has the potential of revolutionizing the entire field of education and social research. It offers a promising learning experience, provides individualized instructions, and enhances education research outcomes in the field of social science. However, considerable attention must be given to the ethical implications, potential biases, and the impact AI has on equity and social justice. It is very instructive to navigate these evolving patterns of learning and understanding to ensure that the use of AI in education and social research synchronizes with the objective of engendering creative thinking, nurturing creativity and providing learners with the required tools to excel in a world where AI has become the major catalyst that drives inventions and innovations.

Future studies should focus on how to identify and mitigate bias in data collection and analysis to ensure that research findings are equitable and do not prolong existing inequalities. Also, scholars should strive to develop methodologies for evaluating the social consequences of AI intervention to ensure that they contribute positively to the academic community. Finally, all stakeholders in this field must commit to the ongoing reflection and reassessment of ethical practices as the technology and societal values evolve.

Further Reading

1. Afzal, A., Khan, S., Daud, S., Ahmad, Z., & Butt, A. (2023). Addressing the digital divide: Access and use of technology in education. Journal of Social Sciences Review, 3(2), 883–895. https://doi.org/10.54183/jssr.v3i2.326
2. Al-kfairy, M., Mustafa, D., Kshetri, N., Insiew, M., Alfandi, O. (2024). Ethical challenges and solutions of generative AI: An interdisciplinary perspective. Informatics, 11, 58. https://doi.org/10.3390/informatics11030058.
3. de Manuel, A., Delgado, J., Parra Jounou, I., Ausín, T., Casacuberta, D., Cruz, M., Guersenzvaig, A., Moyano, C., Rodríguez-Arias, D., Rueda, J., & Puyol, Á. (2023). Ethical assessments and mitigation strategies for biases in AI-systems used during the COVID-19 pandemic. Big Data & Society, 10(1), https://doi.org/10.1177/20539517231179199

4. Jeon, J., Kim, L., & Park, J. (2025). The ethics of generative AI in social science research: A qualitative approach for institutionally grounded AI research ethics. Technology in Society, 81, 102836. http://dx.doi.org/10.2139/ssrn.4784555

JohnBosco Chukwuorji fixing up a research participant in MER Lab at CSU while Dr Yaroslavsky directs him.

Biographies

Dr Godwin Chukwuka Enwu hails from Enugu Ngwo in Enugu State Nigeria. He holds a PhD in Public Administration obtained from the University of Nigeria Nsukka. Currently, he is an Ethnographer and Senior Research Officer at the National Boundary Commission Abuja. He can be reached chubbyshine@gmail.com.

Dr JohnBosco Chika Chukwuorji was a visiting Fulbright scholar at Cleveland State University for his predoctoral research in 2019-2020. A clinical psychologist by training, he is currently a postdoctoral fellow at Michigan State University College of Human Medicine. He also holds a senior lecturer position in psychology at the University of Nigeria. He was a Developing Country Fellow (2022) of the International Society for the Study of Behavioural Development (ISSBD). He can be reached at johnbosco.chukwuorji@unn.edu.ng

Beyond the Hype: Teaching Students to Recognize the Hidden Dangers of AI

Elizabeth A. Radday

Abstract

As generative AI becomes ubiquitous in classrooms and daily life, students need to learn and practice critical thinking. This article explores four of the most prevalent dangers of AI: bias, hallucinations/misinformation, deep fakes, and emotional attachment to chatbots. Educators need to not only be aware of these pitfalls of AI but also need to prepare students to look for and face these dangers with their human intelligence. Ethical, informed AI use begins with awareness, and that awareness must start in K-12 schools.

Keywords: AI Literacy • Digital Media Literacy • AI in K-12 Education.

Artificial Intelligence (AI) has made an indelible mark on the landscape of education. In late 2022, the release of generative AI through platforms such as ChatGPT, Claude, Perplexity, and Gemini immediately impacted middle and high school classrooms as students learned about these Large Language Models (LLMs) that could produce written work within mere seconds. Humanities teachers feared that essay writing would become a lost art. Math teachers empathized as they had been fighting the battle against cheating with the use of PhotoMath and other apps that could solve most homework problems in mere seconds. In higher education, universities and colleges immediately feared how easy it would be to cheat on many traditional assignments that were not done in the classroom under the watchful eye of the professor.

The fears about cheating and student work are not unfounded. Many teachers have spent the last two years thinking of new ways to create durable assessments that cannot be completed by an AI chatbot. These teachers have embraced projects, presentations, and hands-on experiences that could not be done by a LLM. Another, albeit smaller, group of teachers saw AI as an opportunity to elevate the work students could do independently and encouraged students to lean into the capabilities of AI.

In K-12 schools across the United States, teachers are rapidly learning about and adopting new tools with artificial intelligence. Initially teachers looked for ways to make their work more efficient and then turned to ways to use AI to make their lessons and assessments more impactful. At first, teachers tried to take their traditional assignments and prevent students from

cheating. That soon proved to be an exhausting game of cat and mouse, with teachers trying to "catch" students using AI and students getting more sophisticated at using AI without it being detected. Teachers then realized that this was an exhausting game with no winners.

For example, some teachers started by turning back to handwritten assignments completed in class to ensure that students were not using any AI tools. Once teachers were exhausted by these games, they started to explore how it could be helpful in teaching. Many saw that there were ways that students could use AI to improve their writing and allowed students to use approved student-facing AI tools to get feedback on their original written pieces to elevate the final product. Trying to win the war against AI by trying to eliminate all technology from the classroom is unrealistic, as students need to learn how to use these tools to prepare for college and careers.

Now that student facing AI programs that comply with federal and state laws requiring all student data to remain secure and private are available, teachers are slowly extending the use of AI to students in their classrooms in a variety of ways. Teachers have been able to engage students in new ways with content. Middle school students can engage in a conversation with characters from the books they are reading. High school students can debate a bot on topics from climate change to global conflicts without the same anxiety as participating in an oral debate in front of peers. Students can use AI to help brainstorm project ideas, help with time management of projects, and get personalized assistance to study for a test.

The use of generative AI is becoming omnipresent and is now built into so many of the platforms people regularly use, including email and internet searches. Regardless of where one stands on the debate of the value of AI in education, it is out in the world. Students are going to need the skills to use AI ethically, efficiently, and competently to participate in nearly all types of jobs. As middle and high schools prepare their students for higher education and a global workforce, students need to learn how to use AI.

AI RISKS IN EDUCATION

A critical, yet often overlooked, piece of learning how to use AI is learning about the risks in these models. While some people fear that AI may be taking away the need to think critically and will lead to lazier, less intelligent students and adults, many others argue that critical thinking will be *more* important than ever as users evaluate the responses generated by an LLM. The questions teachers should be asking now are: what are the risks of AI, and how can students learn to use their critical human intelligence to ensure the accuracy, reliability, and ethical use of information, especially considering AI-generated bias, hallucinations, and deep fakes? Students also need to

understand the potential dangers of chatbot discussions. Four critical AI risks—bias, hallucination/misinformation, deep fakes, and blurred emotional boundaries—must all be recognized and addressed starting with students in K-12 schools.

AI Bias

LLMs are trained on the massive amount of data that comes from the internet, books, and other written sources. However, much of this data originates from the Western world. This means that an LLM is trained on a huge set of data that leaves out a large swath of information from other parts of the world. As a result, the outputs of an LLM often reflect Eurocentric values, perspectives, and cultural assumptions. Because an LLM can only generate responses based on the data it has seen, regional knowledge and diverse cultural perspectives are often excluded, deepening existing information imbalances.

This can easily be demonstrated by trying one of these two prompts with an LLM. Ask an AI model to tell you a short story about a nurse or prompt an LLM to describe a perfect date night meal. It is highly likely that the story about a nurse will feature a woman, often described as caring and nurturing. More often than not, the patient will be elderly. This is a biased view of nursing and can perpetuate untrue stereotypes. The perfect date night meal will likely be a Western-style dinner that includes a meat dish and an alcoholic beverage (unless the user puts in parameters). This is a highly subjective question, but interestingly, nearly every time, the meals and ambiance are very similar.

Yet many people do not consume alcohol, and traditional foods vary greatly across religions and cultures. Even the concept of "romance" is culturally dependent. What may seem like a simple or neutral prompt can reveal the model's limited cultural scope. The "ideal" meal may not reflect the experiences of people with dietary restrictions, religious considerations, or different cultural understandings of relationships. These examples are just the beginning. The point is clear: training data privileges some voices while marginalizing others.

The task of teachers, given this challenge, is to help students recognize and interrogate bias or stereotypes in the outputs generated by language models. This means not just pointing them out but consistently encouraging students to analyze and improve those outputs. Students should begin by identifying embedded assumptions (e.g., nurses are female, patients are elderly) and exploring ways to challenge or reframe them. In fact, users can prompt the LLM to reflect on the biases in its own responses, bringing to light areas that should be further explored and considered. Teaching students to recognize bias requires fostering critical thinking and helping them understand that bias can appear even in scenarios that may initially seem neutral or objective.

Hallucinations and Misinformation

AI outputs often read like they were written by humans. They flow smoothly, use transitional phrases, and present ideas with clarity and confidence. But behind the scenes, there is no human author. A large language model (LLM) is a mathematical algorithm trained on vast amounts of written language. It does not "think" or "know" anything, it simply predicts the next most likely word, based on patterns it has seen in its training data. There is no one "behind the wheel." An LLM is powered entirely by math, statistics, and code. It takes a user's prompt and generates a response that appears coherent and fluent. But that fluency can be deceiving.

One of the most significant risks of using AI is the potential for hallucinations and misinformation. A hallucination is a term coined by the AI community to describe moments when a model produces information that is entirely made up but often sounds plausible. LLMs can also potentially spread misinformation by getting facts wrong as a result of the information on which they have been trained. Because of their confident tone, LLMs can provide convincing but false answers. Experts can often push an LLM to its limit and expose these hallucinations with just a few detailed prompts. But even middle and high school students can find hallucinations and misinformation when using these models to help with homework!

The danger lies in blindly trusting AI output. For instance, an LLM can suggest a recipe using whatever ingredients you have on hand, but it's not a professional chef, and the result could be unappetizing (or worse). This is an example of a hallucinated recipe. It is made up and may or may not actually be something edible. Students have found misinformation in math, science, and literature assignments, where answers are partially or completely incorrect. Sometimes, students notice the errors. Other times, the AI's fluency makes it hard to tell what is true.

For students, these mistakes can cause serious issues. If they unknowingly absorb misinformation or incorrect procedures, they carry that into future learning. Even more concerning is the issue of hallucinated citations. LLMs have been known to invent articles, authors, and case law out of thin air. This has led not only to students being accused of cheating but also to professionals, like lawyers, submitting fake sources in legal documents. These mistakes do not just risk a grade; they can risk careers.

Teachers, students, or any user of an LLM need to engage in critical thinking and fact-checking when generating any content that relies on truth. Just because what an LLM outputs sounds true does not mean that it is. Another way to say this is that fluency does not equal accuracy. Teachers have long been showing students how to evaluate any resource for its credibility. Using an LLM requires the same kind of media literacy to help identify where the information came from and to verify its accuracy. Just like teachers tell students that one website, article, or source is a starting point, the same is true for LLMs. If the student cannot verify the accuracy of what the LLM writes,

it cannot be relied upon to be accurate. Fact-checking AI should be a habit that students learn from a very young age. This builds academic integrity and strengthens research skills as students look for multiple sources to verify their information. These are skills students will need far beyond their school years.

Deepfakes

Deepfakes are manipulated images, videos, or audio recordings that make it appear as though a real person is doing or saying something they never actually did, often without their knowledge or consent. There have been countless examples of celebrities made to look as though they are in places they have never been or saying things they have never said.

While creating fake content is always unethical, deepfakes become especially dangerous when they harm a person's reputation or sway public opinion. During political elections, for instance, a falsified video of a candidate could be shared widely and influence voters. The societal implications are enormous—and with AI, deepfakes are now incredibly easy and inexpensive to create. With just a low-cost subscription to an AI generation tool, even middle school students can generate fake audio of someone's voice or produce entirely fake but realistic-looking images with a single text prompt.

While it is obviously extremely dangerous when election outcomes can be influenced by deepfakes, what is more important to middle and high school students is when their own reputation is harmed by a deepfake. Now peers can create images that "nudify" others, put classmates in compromising situations, and make someone appear to say things that are harmful or inappropriate. The same tools can be used to target teachers, potentially damaging reputations or leading to serious consequences involving administration or law enforcement. While students have faked content in the past, AI makes the process faster, easier, and more convincing.

Schools, now more than ever, need to teach students how to recognize and question what they see and hear. Deepfakes often contain subtle clues - odd hand shapes, mismatched shadows, inconsistent lighting, and lines that don't match up precisely. Just as students learn to fact-check written content from AI, they must also learn to verify images, video, and audio. This means searching for the original source, considering the reliability of the source, and looking for more evidence that confirms the veracity of the digital media asset. Students must use their critical thinking skills to decide whether it is real or manipulated. This is another lifelong skill that students need to practice in school so that as they become adult consumers of digital media, they can identify what is real or manipulated.

The short film Protect Us from WeProtect Global Alliance dives into the many way students, especially teens, have been harmed (or could be harmed) by AI. Teachers and schools need to have clear policies and protocols to protect students and a consistent and appropriate response to cyberbullying

that has only become easier and more harmful with the rise of AI. Students also need to know that there are safe and trusted adults in their lives to whom they can speak should they find themselves in these embarrassing, damaging, and dangerous situations. So, while teachers need to teach students how to spot deep fakes in media, students also need to understand the consequences of engaging in creating fake images, not just the potential legal ramifications but also the way they emotionally harm their peers.

Blurred Emotional Boundaries

Interactions with AI chatbots are becoming more emotionally realistic every week. What began as simple, text-based exchanges has evolved into lifelike conversations powered by voice models and AI-generated avatars. These tools can now hold natural, fluid conversations using human-like voices, facial expressions, and gestures. For many users, it has become increasingly difficult to tell when they are talking to an algorithm and not a person. Additionally, these bots are designed to be endlessly empathetic, encouraging, and agreeable. Their constant positivity and nonjudgmental responses make them easy to talk to, often more pleasant than real people, which can draw users in and encourage them to spend increasing amounts of time with them. Over the past two years, has blurred the line between genuine human relationships and emotionally convincing simulations. Some individuals have developed strong emotional attachments to AI bots, with cases of users considering ending real-life relationships to pursue a "connection" with a chatbot. Teens have been drawn to these interactions, sometimes isolating themselves from real-world friendships in favor of AI companionship. In one tragic instance, a young person died by suicide, believing he could reunite with his AI companion in another life. These examples illustrate that AI doesn't just influence how we think, it can influence how we feel.

Educators need to help students understand the emotional boundaries between real and simulated relationships. Schools can create opportunities for open conversations about the difference between authentic human connection and AI interaction, emphasizing that bots are designed to simulate empathy, not experience it. Teachers should incorporate media literacy and digital ethics into discussions about AI, helping students develop awareness of how easily these systems can imitate emotional intimacy—and why it matters to stay grounded in real-life relationships.

AI Literacy in K-12 Schools

It seems obvious that AI Literacy is now a critical skill that schools need to teach students starting as early as kindergarten. However, there are several challenges with this demand. There is no consensus, yet, on what defines AI literacy and what it means to teach these skills. Additionally, because these skills need to be integrated across the curriculum, it is unclear who should teach which skills and when. When "everyone" owns an initiative, no one owns

the initiative. AI literacy requires a full system collaborative effort to form a coordinated and cohesive plan for students to learn all the necessary skills. AI literacy needs to include all the dangers (bias, hallucination/misinformation, deep fakes, blurred emotional boundaries) along with all the ways that AI can be helpful and useful in education. Students need to learn how to use a variety of tools, while also understanding the ethical implications. This is a huge task that requires more than just a few lessons in one grade or one class. Finally, because AI technology is moving so fast, what students need to know and be aware of is always changing so the curriculum constantly needs revision.

In May 2025 the OECD and European Commission released a draft version of an AI Literacy Framework. This framework offers 22 competencies across the four major domains of AI: engage with AI, create with AI, manage AI, and design AI. This framework is going to have a tangible impact on AI education globally. The draft framework is open for review and feedback until the end of August 2025, and the finalized version will be available in 2026.

Conclusion

The challenges posed by artificial intelligence—bias, hallucinations, deep fakes, and emotional manipulation—are not confined to a single country, culture, or educational system. They are universal issues that impact how people across the globe access, evaluate, and engage with information. As AI continues to evolve at an unprecedented pace, educators everywhere have a shared responsibility to ensure that students can navigate this technology confidently, ethically, and critically.

During my Fulbright in Finland, a country known for its deep commitment to equity and student well-being, I saw firsthand how thoughtfully designed education systems can prepare students for complex societal challenges. Finland prides itself on its innovation in technology. Nokia, the world's leading cell phone manufacturer in the late 1990s and early 2000s, is a Finnish company. While I was in Oulu, the home city of Nokia, I was able to witness the "Polar Pitch," an event in February where new small businesses pitched their startups to a panel of judges and investors. The catch was that they could pitch their product only for as long as they could stand in the icy water! This spirit of innovation and future focused thinking about technology was evident in many ways throughout my time in Finland. I spent a lot of time visiting vocational schools. Students were encouraged to start their own small businesses and use their skills to earn money before graduating. Many vocational schools had small stores where students could sell their products. Students were allowed to use tools and machines to manufacture goods for sale. I was impressed with the entrepreneurial spirit of Finns and their pride in technology. Angry Birds, one of the first viral mobile games, was invented

by a Finn. Finns were always happy to share about these Finnish innovations, constantly eager to remind visitors that even though they are a small country many people haven't even heard of, they are not to be forgotten about in the technology race.

It was no surprise to me that as generative AI became widely available to the public, Finland would want to be seen as a leader in this space. Finland made news in the US in 2023 because they ranked #1 of 41 European countries in resilience against misinformation (Gross, 2023). The United States has reported on how Finland has started AI literacy across all grades and is taking note on how to bring that same literacy to students in the US. Starting by exposing bias, misinformation and hallucinations to students, is one way to do that. Whether in Helsinki, Connecticut, or anywhere else around the world, classrooms must now serve as the first line of defense against misinformation and manipulation. Teaching young people to question, verify, and reflect is no longer just good pedagogy, it is an essential part of preparing them to be thoughtful global citizens in a world increasingly shaped by artificial intelligence.

"Teaching young people to question, verify, and reflect is no longer just good pedagogy, it is an essential part of preparing them to be thoughtful global citizens in a world increasingly shaped by artificial intelligence."

Further Reading

1. Gross, J. (2023, January 10). How Finland Is Teaching a Generation to Spot Misinformation. New York Times. https://www.nytimes.com/2023/01/10/world/europe/finland-misinformation-classes.html
2. OECD (2025). Empowering learners for the age of AI: An AI literacy framework for primary and secondary education (Review draft). OECD. Paris. https://ailiteracyframework.org
3. European Commission: Directorate-General for Education, Youth, Sport and Culture, *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*, Publications Office of the European Union, 2022, https://data.europa.eu/doi/10.2766/153756

Elizabeth and her family at the SnowCastle of Kemi, Finland in 2016.

BIOGRAPHY

Elizabeth Agro Radday, Ed.D., has worked in the education field for 25 years and was a Fulbright Distinguished Awards in Teaching Grantee to Finland in 2016. She is currently the Director of Research and Innovation at EdAdvance, a Regional Education Service Center in Connecticut, and the co-host of the popular podcast ChatEDU: A podcast about AI in Education. She is passionate about teaching students how to use technology and is a strong advocate for students doing Personal Interest Projects to learn real-world skills by exploring their passions. She has presented nationally and internationally on AI in K-12 schools and on Capstone/Personal Interest Projects. Her book, *Learning They'll Love*, will be published in November 2025 with ASCD.

Harnessing AI for Information Diplomacy: Addressing Algorithmic Bias in Cross-Cultural Service-Learning

Bethany S. McGowan

Abstract

As artificial intelligence (AI) transforms education, its integration into service-learning presents opportunities and challenges for algorithmic literacy, infodemic management, and global engagement. This article examines how AI-enhanced service-learning fosters analytical and ethical skills, drawing from a Fulbright-sponsored course in Finland and a Diplomacy Lab project. It explores AI's pedagogical implications, role in information diplomacy, and faculty development, emphasizing AI's potential to reshape service-learning while advancing critical thinking and algorithmic literacy.

Keywords: Artificial Intelligence • Service-Learning • Information Diplomacy • Infodemic Management • Experiential Education.

My exploration of artificial intelligence (AI) and information diplomacy began with an early career focus on policy-driven information management. As a postgraduate fellow with the Congressional Research Service, I examined how research supports public decision-making, particularly in areas where information governance intersects with public trust. This foundation was further developed during my internship with the U.S. Department of State's Humanitarian Information Unit. There, I applied my skills in information systems and data analysis to support conflict prevention and humanitarian response efforts. These experiences introduced me to the potential of data-driven insights to inform strategic policy discussions. They also reinforced my long-standing commitment to the intersection of technology, policy, and global information ecosystems, a focus that continues to guide my work today.

As an Associate Professor in Information Studies, I explore how emerging technologies influence public information systems, with implications for access, trust, and policy. My teaching and research explore the role of artificial intelligence in shaping the creation, circulation, and governance of public information. In 2022, I co-led a Diplomacy Lab project – Strategies for Identifying Mis/Dis/Malinformation – which engaged students in using AI-driven social listening tools to monitor mis/dis/malinformation trends on open social media platforms. Diplomacy Lab, a U.S. Department of State initiative, connects university faculty and students with real-world policy challenges to generate actionable insights for diplomatic priorities.

Through the project, students designed and prototyped machine learning models capable of real-time data analysis, gaining firsthand experience with the complexities of digital mis/dis/malinformation. The experience served not only as a case study in applied AI, but also as a turning point in how I conceptualize teaching. It demonstrated that integrating AI into service-learning requires a pedagogical shift from traditional problem-based learning models to AI-assisted, data-driven approaches that foreground critical interpretation, ethical reasoning, and civic engagement.

The opportunity to expand this work internationally through the Fulbright U.S. Scholar Program provided a critical next step. Finland, a leader in mis/disinformation resilience and digital literacy education, was an ideal setting to examine how AI-enhanced approaches can contribute to global mis/disinformation management strategies while also redefining instructional methodologies. While teaching a graduate course--Infodemic Management for Diverse Communities--I introduced students to AI-supported mis/disinformation tracking systems, namely the World Health Organization (WHO) Early AI-supported Response with Social Listening platform (EARS), and explored the ethical, social, and policy implications of AI-assisted information interventions.

Beyond pedagogy, this experience deepened my engagement with AI's role in information diplomacy--a strategic practice that, as Emily Metzgar suggests in *On the Policy and Politics of U.S. Information Diplomacy* (2021), refers to the use of information--particularly accurate, credible, and strategically curated content--as a soft-power tool in international engagement. Drawing from Metzgar's commentary, it underscores the need for intentional, ethically guided communication to foster understanding and counter mis/disinformation.

In this context, AI becomes more than a technical tool—it functions as an instrument of civic engagement and digital diplomacy. Embedding AI into instructional spaces reveals its potential to prompt critical reflection on how information is framed, filtered, and acted upon. To navigate this complexity, students and educators alike must cultivate AI literacy, which Long and Magerko define in their 2020 ACM Conference on Human Factors in Computing Systems talk titled *What is AI Literacy? Competencies and Design Considerations* as: "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace." These competencies are essential for addressing algorithmic bias, ethical ambiguity, and the sociopolitical implications of automated systems.

AI AND INFORMATION DIPLOMACY IN SERVICE-LEARNING

AI is reshaping the landscape of service-learning by expanding how students engage with information diplomacy in practice--applying digital tools to navigate complex global issues, analyze public discourse, and collaborate across cultural contexts. Integrating AI tools into academic instruction and community-based engagement enables students to move beyond theoretical data analysis and directly engage with real-world information challenges. In her 2024 study *AI-Facilitated Critical Thinking in an Undergraduate Project-Based Service-Learning Cours*e, Sara Kimmel demonstrates how incorporating AI into service-learning strengthens critical thinking and reflective skills by engaging students in iterative analyses of complex social issues. Similarly, in their 2024 study, *Leveraging Artificial Intelligence in Project-Based Service Learning to Advance Sustainable Development: A Pedagogical Approach for Marketing Education*, C. M. Dubay and Melanie Richards present the AISLE model, illustrating how ethically integrated AI in project-based service-learning can support student learning outcomes while advancing the United Nations Sustainable Development Goals. Both studies' findings suggest that the integration of AI elevates service-learning by positioning students as active co-creators of knowledge, capable of leveraging digital tools to address complex social issues, communicate across cultures, and navigate the ethical terrain of emerging technologies.

I found similar results in my own work. Building on my Diplomacy Lab course, my Fulbright experience at Tampere University expanded my AI-informed service-learning approach to a global learning environment. I designed and taught a graduate course--Infodemic Management for Diverse Communities--that encouraged students to explore how AI could be applied to address mis/dis/malinformation challenges in multicultural settings, with a particular emphasis on the intersection of AI, public health, and community engagement. This model required students to interpret AI-generated insights and to assess AI's limitations, critically questioning how algorithmic bias, data transparency, and contextual inaccuracies shape digital narratives.

AI-driven tools, such as the WHO EARS, provided a dynamic framework for students to examine mis/dis/malinformation patterns in real time. By engaging with AI tools, students were able to identify trends, assess the credibility of online discourse, and develop targeted interventions to address misinformation within different sociocultural contexts. Yet this process also revealed the complexities of AI as an instructional tool, requiring constant human oversight, cultural adaptation, and ethical evaluation. Students encountered instances where the underlying data models failed to adequately capture the nuance of underrepresented languages or cultural contexts, leading to misclassifications or gaps in interpretation. These limitations were especially evident in multilingual or informal online discourse, where meaning

was shaped by local idioms, humor, or shifting sociopolitical references. AI alone could not provide fully actionable insights; students had to apply critical thinking to refine its outputs, demonstrating that AI must function as an assistive rather than authoritative tool in service-learning.

These cross-cultural differences significantly shaped how AI tools were integrated and interpreted across the two course settings. In Finland, digital literacy is introduced as early as age six and reinforced through a nationally coordinated approach involving schools, libraries, and NGOs. Students arrive in university with well-developed critical thinking skills and a deeply engrained understanding of how to evaluate digital content. This sustained investment in digital literacy is widely recognized as a key contributor to Finland's consistent top rankings in European media literacy indexes and it informed the high level of critical engagement I observed among Finnish students using AI tools in my Fulbright course. They questioned not only the reliability of AI-generated outputs, but also the governance models behind the data infrastructure itself, often framing their critiques within public health and policy ethics.

By contrast, the upper-class undergraduates in my U.S.-based Diplomacy Lab course often required more scaffolding to critically assess AI outputs. While the undergraduate versus graduate level difference between the two courses makes a direct comparison impossible, a 2021 Stanford study, *National Study of High School Students' Digital Skills Paints Worrying Portrait*, offers some insight. It found that only 0.1% of 3,446 U.S. high school students correctly identified the source of a misleading video and just 4% recognized a fossil fuel–funded climate change website as biased. The researchers concluded that the findings presented "a concerning picture of American students' ability to figure out who produced a given story, what their biases might have been, and whether the information is reliable." These results underscore a systemic need for AI literacy education in the U.S. and highlight the instructional demands of introducing AI tools into American service-learning environments.

Across both U.S. and Finnish contexts, service-learning created space for students to see AI not as an objective authority, but as a socially embedded, ethically contingent tool. Students actively collaborated with community health organizations, global health NGOs, and policymakers to apply AI-driven insights to real-world problems. Both American and Finnish students brought diverse perspectives on institutional trust, media ecosystems, and misinformation resilience, allowing for comparative discussions on how AI tools could be adapted to different information environments. Yet, the stark contrast in baseline AI literacy and institutional trust revealed that AI-supported instruction must be carefully adapted to national educational cultures, rather than applied uniformly across global classrooms.

By incorporating AI into service-learning, students moved beyond passive analysis to action-based learning, where they applied AI-driven findings to design public communication strategies, develop educational interventions, and contribute to community-centered mis/dis/malinformation mitigation efforts. This experience reframed students as active participants in AI-

enhanced information diplomacy, equipping them with the skills to leverage AI responsibly in civic and global engagement efforts. More importantly, it prepared them to think critically about AI as a tool for social impact, developing adaptive AI literacy skills that will be essential as AI becomes further embedded in evidence-based decision making.

> ***"By incorporating AI into service-learning, students move beyond passive analysis to action-based learning, where they applied AI-driven findings to design public communication strategies and contribute to community-centered mis/dis/malinformation mitigation efforts."***

Faculty Development, AI Integration, and the Challenges of AI in Service-Learning

While much of the discussion around global service-learning focuses on student outcomes, the integration of AI-driven tools into experiential learning environments presents unique challenges for faculty professional development. Exchange programs such as the Fulbright U.S. Scholar Program and public-private federal partnerships like Diplomacy Lab have expanded my academic growth, instructional strategies, and research collaborations, but they have also revealed barriers to long-term AI integration in education. Faculty must navigate the technical, ethical, and sustainability challenges of embedding AI into instructional models while ensuring that AI remains a tool for enhancing learning rather than replacing critical engagement and human interpretation.

One of the primary challenges of incorporating AI into service-learning is navigating the balance between technological advancement and ethical responsibility. During my instruction, I found that AI provided valuable insights into public discourse trends and misinformation patterns. For example, the WHO EARS tool proved especially useful for identifying trending public health topics during Fall 2023. While COVID-19 and the COVID-19 vaccine remained part of the conversation, my students and I also surfaced concerns about the severity of the seasonal flu and broader debates around public health preparedness. EARS enabled us to track how these topics evolved in near real time in Finland. They then used these insights to guide their selection of community partners and align their project focus with the issues most actively discussed in the public sphere, such as partnering with local health agencies to address flu communication or with NGOs focused on vaccine equity.

However, the process also revealed the limitations of AI in cross-cultural contexts. The EARS system at times struggled with accuracy, contextual understanding, and bias mitigation. For instance, nuances in spoken Finnish and casual written language, common in online discourse—differ significantly from the formal Finnish used to train many AI language models. This occasionally led to the misclassification of ironic or idiomatic posts,

prompting students to question the reliability of algorithmic interpretations and reinforcing the need for human oversight in AI-supported analysis. Similarly, in the U.S.-based Diplomacy Lab course, students attempted to build their own social listening tool using open-source platforms and publicly available data. While the exercise fostered hands-on learning about data preprocessing and machine learning, it also revealed how access to high-quality, representative training data, particularly for marginalized communities or niche issue areas, was limited or unevenly distributed. These challenges highlighted that AI's instructional value is inseparable from broader issues of data equity, cultural specificity, and infrastructure access.

These limitations point to a broader gap in faculty training: AI literacy is not yet a standard component of professional development, particularly for instructors integrating experiential or service-learning models. As Hervieux and Wheatley observe in their 2024 white paper, Building an *AI Literacy Framework: Perspectives from Instruction Librarians and Current Information Literacy Tools*, librarians and instructors are increasingly engaging with generative AI in the classroom, yet few have access to structured frameworks that address its ethical, societal, and technical dimensions. Most continue to rely on the ACRL Framework for Information Literacy in Higher Education, which offers a strong foundation for critical inquiry but lacks guidance specific to the pedagogical challenges posed by AI.

As a result, instructors must take on dual roles as both educators and AI facilitators, guiding students not only on how to use these tools, but also on how to critically evaluate AI outputs, recognize algorithmic bias, and navigate ethical considerations across cultural and disciplinary contexts. Without such guidance, students risk interpreting AI-generated insights as objective truths rather than socially constructed outputs that require thoughtful interpretation and contextual judgment.

Addressing this gap will require targeted investment in faculty-facing AI literacy development. Universities must prioritize interdisciplinary workshops, team-teaching models, and course release opportunities that allow faculty to explore AI's evolving role in applied learning. This includes building shared frameworks for ethical evaluation, bias detection, and curricular integration. Faculty must ultimately move from being occasional users of AI to becoming AI pedagogical designers, ensuring that these tools are not applied as generic, one-size-fits-all solutions, but are thoughtfully adapted to the civic, cultural, and disciplinary contexts in which students learn and engage.

Another critical limitation of AI in service-learning is its accessibility, usability, and long-term sustainability. While AI tools offer data-driven insights, they often require technical expertise that can present barriers for students, instructors, and community partners. Many platforms lack user-friendly interfaces or culturally adaptive features, complicating their integration into community-based projects. Resource constraints such as limited institutional access to AI software, training opportunities, and computational infrastructure further hinder the scalability of AI-enhanced service-learning initiatives.

For example, while using the WHO EARS platform in my Finnish course, students encountered challenges related to the tool's limited regional customization, which reduced its effectiveness in capturing mis/dis/malinformation trends specific to Finland. This required substantial human interpretation and local expertise to supplement the AI-generated outputs. Compounding these issues, the EARS platform is no longer publicly available due to broader budget constraints at the World Health Organization, raising concerns about the reliability and longevity of open, external AI tools in academic settings. These limitations underscore the need for institutional investment not only in training and infrastructure, but also in the long-term integration, support, and adaptability of AI tools. Without such commitment, AI risks remaining a series of isolated pilot efforts rather than becoming a sustainable and embedded element of service-learning pedagogy.

Another challenge is faculty preparedness in guiding students through AI ethics. Algorithmic bias, data privacy concerns, discussions related to soft v. sharp power, and mis/dis/malinformation amplification risks require careful consideration when using AI as a tool for community engagement and public diplomacy. These issues became particularly relevant in my work despite AI's capabilities, human oversight was essential in both courses to contextualize AI-generated insights, prevent overreliance on automated decision-making, and ensure that interventions aligned with ethical communication principles. Yet, many instructors lack formal training in AI ethics—faculty development programs must evolve to include best practices for addressing bias, transparency, and mis/dis/malinformation risks in AI-assisted education.

In addition to institutional support and ethics training, exchange programs can play a crucial role in helping faculty adapt to AI-enhanced instructional models. The Fulbright Program enabled me to collaborate with European scholars on AI-driven service learning, expanding my global research network. Through an Inter-Country Travel Award, I visited Sweden, where discussions on AI's role in public trust, digital ethics, and mi/dis/malsinformation policy further enriched my perspective. These international engagements reinforced the need for multidisciplinary, multinational approaches to AI integration in education—ones that combine technological expertise with cultural and ethical sensitivity.

As AI continues to shape service-learning and information diplomacy, faculty must be equipped with critical digital literacy, ethical AI practices, and pedagogical adaptability. Addressing the limitations of AI in instructional settings requires more than individual effort, it demands sustained institutional investment in both experiential education and AI literacy infrastructure. My experiences have shown that service-learning flourishes when universities prioritize cross-sector collaboration and civic engagement. For example, my university supports a dedicated center for experiential education, a robust

public policy institute that fosters community partnerships, and a strong Fulbright advising program—each of which played a significant role in helping me scaffold civic learning projects where AI tools can be meaningfully applied.

Similarly, my host institution in Finland, Tampere University, demonstrated strong institutional support through its WHO Collaborating Centre for Health in All Policies and the Social Determinants of Health, which provided public health expertise and global policy context for student projects. My host department also had a long-standing history of hosting and supporting Fulbright Scholars, offering both institutional continuity and a globally engaged academic culture that made service-learning with AI both feasible and impactful.

Final Reflections

Reflecting on my Fulbright journey, I am struck by the transformative potential of AI in global service-learning. My experience integrating AI tools into a service-learning course in Finland revealed that AI is not merely a tool for data analysis, but a powerful catalyst for global engagement, digital literacy, and policy-informed education. At the same time, my experiences underscored the challenges of incorporating AI into service-learning on a global scale. AI tools are not universally accessible or equally effective across cultural and linguistic contexts, requiring human oversight, adaptability, and interdisciplinary collaboration. The sustainability of AI-enhanced service-learning also depends on ongoing faculty training, institutional support, and cross-sector partnerships, ensuring that AI remains a meaningful complement to instruction rather than a short-term innovation.

Moving forward, I am committed to further integrating AI into service-learning frameworks globally. In Summer 2024, I completed a cultural residency program in France that explored the policy dimensions of equitable information access. My work focused on how the placement of cultural institutions—such as libraries—and the design of their services can maximize their public impact, particularly in relation to the challenges of navigating information in an AI-driven world. That experience reinforced the importance of interdisciplinary and international collaboration in shaping the next generation of AI-enhanced service-learning models. My future research and instructional development will focus on how AI can be more effectively integrated into experiential learning in libraries while addressing equity, accessibility, and long-term sustainability.

My Fulbright experience reaffirmed my belief that education must go beyond transmitting knowledge; it must be a catalyst for social change. By integrating service-learning, AI, and information diplomacy, I will continue to foster equity, trust-building, and community empowerment, ultimately contributing to a more informed, ethically responsible, and globally connected society.

Further Reading

1. Artificial intelligence (AI) is increasingly shaping global information ecosystems, raising critical concerns about algorithmic bias, misinformation management, and ethical governance. For a detailed overview of AI's role in combating misleading information, see The Oversight Board's (2024) Content Moderation in a New Era for AI and Automation. Available at: https://www.oversightboard.com/wp-content/uploads/2024/09/Oversight-Board-Content-Moderation-in-a-New-Era-for-AI-and-Automation-September-2024.pdf
2. Although the EARS (Early AI-supported Response with Social Listening) tool described is no longer available, it served as a foundational component of the service-learning approach. The World Health Organization (WHO) continues to support a broader suite of AI-enhanced tools and strategies for infodemic management. To explore WHO's current global initiatives in this area, visit: https://www.who.int/health-topics/infodemic
3. Service-learning provides a critical platform for students to engage with AI in real-world problem-solving, particularly in culturally diverse and justice-oriented contexts. For an in-depth discussion of AI's role in cultural competency, service-learning, and expanding access to justice, see Robinson, G. (2025). *Artificial Intelligence: Cultural Competency, Service Learning, and Community Service Expanding Access to Justice. Cambridge University Press. ISBN: 979-8-8230-4178-2.*

As a Fulbright U.S. Scholar in Finland, Bethany McGowan conducted research and teaching at Tampere University, pictured here on campus in Tampere.

BIOGRAPHY

Bethany McGowan is an Associate Professor in the Libraries and School of Information Studies at Purdue University, specializing in service-learning, information diplomacy, and health information literacy. As a Fulbright Scholar to Finland (2023–2024), McGowan taught and conducted research on global health and information challenges, emphasizing the roles of information diplomacy and infodemic management in fostering cross-cultural understanding. Her work integrates academic instruction with public diplomacy frameworks to address global information challenges and promote community resilience. She can be reached at bmcgowa@purdue.edu.

AI ETHICS AND GOVERNANCE: TOWARDS THE IMPLEMENTATION OF THE HUMAN-CENTERED AI (HCAI) FRAMEWORK

ALIZA D. RACELIS

ABSTRACT

Artificial intelligence (AI) has profoundly transformed our lives. Concerns about AI's evolution and potential dangers have led to the development of regulations in various regions. While the regulatory approaches differ, they share a goal: ensuring AI benefits society while minimizing negative impacts. This paper argues that these regulations should incorporate the principles from the Human-Centered AI framework. This is a shortened version of a paper delivered at the *Northern Philippines Business Research Conference* in February 2025.

Keywords: Artificial Intelligence • Governance • Human-centered AI •

Artificial intelligence (AI) has profoundly changed and will continue to change our lives. AI is being applied in an increasing number of fields and scenarios, such as autonomous driving, medical care, media, finance, industrial robots, and internet services. The widespread application of AI and its deep integration with the economy and society have improved efficiency and produced benefits. At the same time, it will inevitably impact the existing social order and raise ethical concerns. Ethical issues, such as privacy breaches, discrimination, unemployment, and security risks brought about by AI systems, have generated significant concern among individuals. Consequently, AI ethics has become an important research topic in academia as well as a topic of common concern for individuals, organizations, societies, and countries.

Recent developments in AI have generated significant interest from media and the general public. As AI systems (e.g., robots, chatbots, avatars, and other intelligent agents) evolve from being perceived as tools to being perceived as autonomous agents and teammates, an important focus of research and development is understanding the ethical impact of these systems. Critical questions have emerged: What does it mean for an AI system to make a decision? What are the moral, societal, and legal consequences of their actions and decisions? Can an AI system be held accountable for its actions? How can these systems be controlled once their learning capabilities bring them into

states that are possibly only remotely linked to their initial, designed setup? Should such autonomous innovation in commercial systems even be allowed, and how should AI use and development be regulated? These and many other related questions are currently the focus of much attention.

The Philippines Context

In the Philippines, the challenges posed by AI must be understood within the broader context of business ethics concerns identified by educators and practitioners—particularly the country's widespread poverty that persists despite impressive economic growth. Business ethics professors and practitioners are encouraged to ensure the ethical use of AI in business operations, especially how algorithms can produce outcomes that lead to unintended consequences, such as discrimination, job displacement, privacy violations, and other societal impacts. The Philippines is steadfast in achieving the 17 Sustainable Development Goals (SDGs) by 2030. AI-based research initiatives of its Department of Science and Technology (DOST) are anchored in these goals to achieve better healthcare, economic growth, clean energy, smart cities, smart farming, and climate change mitigation. However, the Philippines' House Bill No. 7913 primarily focuses on establishing the Philippine Artificial Intelligence Council and the AI Research and Development Program. While the bill acknowledges the potential benefits of AI, it falls short in explicitly addressing crucial aspects such as the ethical implications, fairness, potential biases, and the societal impact of AI technologies.

Global AI Ethics Standards and Frameworks

The adaptation of principles and concepts for AI ethics should be based on internationally recognized standards. In November 2021, UNESCO adopted the *Recommendation on the Ethics of Artificial Intelligence*, marking a significant milestone in developing global standards for AI ethics. Supported by all 193 member states, this recommendation serves as a normative framework to address ethical concerns related to AI and to foster trustworthiness throughout the AI system life cycle. It places transparency, fairness, and the protection of human rights and dignity at its core. Along with these, the Center for AI and Digital Policy (CAIDP) emphasizes addressing the connection between AI and human rights. CAIDP, a non-profit organization, is committed to ensuring that advancements in AI contribute to a more equitable and fair society. It advocates for a world where technological advancements are made in hrmony with respect for human rights, rule of law, and democratic institutions.

Like the European Union's General Data Protection Regulation (GDPR) in 2018, the EU AI Act could become a global standard, determining to what extent AI has a positive rather than negative effect on life. The EU's AI regulation is already making waves internationally. In late September

2024, Brazil's Congress passed a bill creating a legal framework for artificial intelligence. There are, however, several loopholes and exceptions in the EU law. These shortcomings limit the Act's ability to ensure that AI remains a force for good.

UNESCO's Recommendations on the Ethics of AI appear to present the most robust AI guidelines among the global guidelines. Their recommendations have set the standard and served as a benchmark for developing other AI guidelines. They recommend adapting principles for an ethical framework that promotes responsible development and use of AI technologies. UNESCO's guidelines emphasize the importance of human rights, transparency, explainability, and accountability in AI systems.

Navigating the Regulatory Landscape

We can define regulation of AI broadly as including not only legislation and government policies but also professional norms and technical standards. Central to this task is the question, What parameters are required? Although national and international government bodies play a defining role here, other players are also influential. Defining rules for something as extensive, complex, and versatile as a system technology brings numerous challenges, problems, and dilemmas. One of the best known is the "Collingridge dilemma." On the one hand, a new technology is difficult to regulate in the early phase because much remains unclear regarding its workings and effect. Moreover, the need for regulation is initially less apparent. Later, once the technology's effects on society are more conspicuous, it becomes clear what regulation is needed and why. By then, however, many of the decisions taken earlier are difficult to reverse. A further complication is that power structures develop around a technology, and these cannot be modified easily or quickly. The Collingridge dilemma is exemplified by the architecture of the internet, which was developed in a spirit of openness and market freedom. Today it is clear that many safety and security issues were not adequately addressed by the original design. Rectification of these design flaws at this stage would require large sections of the internet to be completely restructured. .

The Current State of AI Governance

Embedding or integrating AI into society depends on the existence of frameworks, and therefore regulation. Now that the technology is making the transition from the lab to society, its effects on the economy and society are subject to widespread scrutiny. This has led to debate about the nature of the regulatory measures needed to ensure that AI is properly integrated in society and government processes. Attention has focused not only on the opportunities but also on AI's potential negative consequences. Hundreds of guidelines, codes of conduct, private standards, public-private partnership models and certification schemes have been developed with a view to both promoting opportunities and addressing adverse repercussions. One of the

more important initiatives is the *European Union's AI Act*. Many existing legal provisions and frameworks are potentially applicable to AI, ranging from fundamental rights to liability law, intellectual property rights and the rules on archiving and evidence. In other words, the effects of AI are now controlled through a wide range of frameworks and specific rules, many more of which are likely to be laid down in the years ahead.

Distinguishing Human Intelligence from Artificial Intelligence

Given these challenges, a new ethic of technological development, based on the unconditional priority of public interest and security of the individual, ought to be developed. A critical distinction must be made: the distinction between human intelligence and "artificial intelligence." According to Turing, artificial intelligence mimics humans in the process of preparing and making decisions. This kind of intelligence is very useful in organizational activities, as it offers opportunities to improve human performance by extracting relevant information from large datasets and by predicting unexpected events, by doing so in a fraction of the time it takes humans to do it. Through its imitative abilities, AI is able to identify information patterns that optimize work-related trends. However, humans possess cognitive abilities that represent true intelligence – human intelligence. Being in an open system, humans must respond accordingly to exogenous influences. This mode requires a creative approach to the formation of future strategy, manifested in the ability to correctly respond to sudden changes in the situation and to anticipate the possible developments, as well as to correctly perceive distorted information. All this requires a rational and radical concept of "responsibility."

The Race for Trustworthy AI

It has been argued that a race to AI regulation ought to be pursued, with ever-louder calls being made for regulators to look beyond the benefits and ensure that AI is trustworthy – that is, legal, ethical, and robust. Besides minimizing risks, such regulation could facilitate AI's uptake, boost legal certainty, and also contribute to advancing countries' positions in the race. Indeed, a new playground for global regulatory competition seems to be emerging, which in the best-case scenario pushes governments—amid uncertainty as to the technology's impact, the impact of regulatory intervention, and the cost of non-intervention—to find the most appropriate balance between protection and innovation. By striving for such balance in their own distinct manners, countries can compete through regulation to attract those ingredients that render them a competitive force on the global AI market, while exploring the best approaches to protect their citizens.

The Rome Call and Interfaith Perspectives

The Rome Call for AI Ethics (www.romecall.org), finalized in February 2020, committed signatories to follow principals of transparency, inclusion, accountability, impartiality, reliability, security, and privacy. Religious faiths have played and will continue to play a role in shaping a world in which human beings are at the center of the concept of development. It was argued at the February 2020 event that the ethical development of AI must be approached from an interfaith perspective. In the face of radical transformations that digital and intelligent technologies are producing in society, the three Abrahamic religions together provide guidance for humanity's search for meaning in this new era.

The Emergence of Human-Centered AI

While the technology-centered approach has dominated the development of AI technology, researchers have individually explored a range of human-centered approaches to address the unique issues introduced by AI technology. These include humanistic design research, participatory design, inclusive design, interaction design, human-centered computing, and social responsibility. To respond to AI ethical challenges, Stanford University established a Human-Centered AI (HCAI) research institution, focusing on ethically aligned design. HCAI suggests strategies that support human self-efficacy, creativity, responsibility, and social connections. Researchers, developers, business leaders, policymakers, and others are expanding the technology-centered scope of artificial intelligence (AI) to include HCAI ways of thinking. This expansion from an algorithm-focused view to embrace a human-centered perspective can shape the future of technology to better serve humanity.

Implementing the HCAI Framework

I recommend the Human-Centered Artificial Intelligence (HCAI) framework for designing and assessing AI systems and tools. HCAI clarifies how to (1) design for high levels of human control and high levels of computer automation so as to increase human performance, (2) understand the situations in which full human control or full computer control is necessary, and (3) avoid the dangers of excessive human control or excessive computer control. Achieving these goals will support human self-efficacy, creativity, responsibility, and social connections. In summary, AI ought to amplify, augment, enhance, and empower people. Educators, designers, software engineers, product managers, evaluators, and government agency staffers can build on AI-driven technologies to design products and services

> ***"A greater emphasis on human-centered AI will reduce fears of AI's existential threats and increase benefits for users and society in business, education, healthcare, environmental preservation, and community safety."***

that make life better for users, enabling people to care for each other. A greater emphasis on Human-Centered AI will reduce fears of AI's existential threats and increase benefits for users and society in business, education, healthcare, environmental preservation, and community safety.

Further Reading

1. Alfiani, F. R. N. & Santiago, F. (2024). A comparative analysis of artificial intelligence regulatory law in Asia, Europe, and America. *SHS Web of Conferences*. https://doi.org/10.1051/shsconf/202420407006
2. European Union (2024). *Artificial intelligence act (Regulation (EU) 2024/1689)*. https://artificialintelligenceact.eu/the-act/
3. Rosales, M. A.; Magsumbol, J. V.; Palconit, M. G. B.; Culaba, A. B.; & Dadios, E. P. (2020). Artificial intelligence: The technology adoption and impact in the Philippines. *27th International Computer Science and Engineering Conference (ICSEC)*. https://doi.org/10.1109/ICSEC59635.2023.10329756
4. Shneiderman, B. (2022). *Human-centered AI. United Kingdom: Oxford University Press.*

Aliza Racelis gifting Executive Director Dr. Julio Amador (Fulbright Philippines) with a copy of her textbook Business Ethics and Social Responsibility, December 2022.

Biography

Aliza Racelis received her PhD in Business Administration from the University of the Philippines in April 2010. She is a Management and Business Ethics professor at the University of the Philippines Business School. Her current research interests are in the areas of Business Ethics, Corporate Governance, Virtue Theory, Social Responsibility, Transcendental Leadership, and Sustainability. In 2018, she was granted a Fulbright scholarship to do research and teaching in the U.S. during the Fall Term.

Artificial Intelligence and Equity in Education: The Way Ahead

Aldan Creo

Abstract

How can we use AI to promote equity in education across cultural and socioeconomic divides? To answer this question, this article explores the old and new challenges that AI presents, including cost, implementation disparities, the closed nature of industry, inherent biases, and the impracticability of punitive approaches to AI use. I also discuss the way forward and how we can work to make AI a force for equity in education.

Keywords: Artificial Intelligence • Education • Equity • Bias • AI-Generated Text Detection

The Promise and the Reality

Artificial Intelligence is a disruptive technology that promises to revolutionize the academic landscape. If we use it properly, we can promote equity at all levels of education.

Sounds cliché? Rightly so, we have been here before. I could have written almost the same line in the early 2000s. The internet was supposed to democratize access to information and make quality education a reality for everyone, regardless of location or socioeconomic status. While some progress has undoubtedly been made, inequity unfortunately remains a reality.

In this article, I hope to provide some insights into the challenges and opportunities that AI presents in the context of education, with a particular focus on the pathways to ensuring that AI helps to promote equity rather than exacerbate existing inequalities, and how Fulbrighters are uniquely positioned to play a crucial role in this transformation.

Old Challenges That Persist

Several challenges that have persisted since the early days of the digital age continue to impede the realization of the promise of technology in education. Although AI was first conceptualized around the 1960s, its practical applications in education have only recently emerged, thanks to large language models. But first, we should define "equity" in this context. Here, I'll use the term to mean that all students, regardless of background, have the same opportunities to succeed in school and in life, sometimes referred

to with the (similar but not identical) idea of "equality." There are two sides to the intersection of equity and education systems: equal access to quality education and education in values shaped by equity. I believe both are equally important for societal progress, and AI can help us achieve both.

However, one reason often cited for why new technologies have not supported equity: their cost to less-developed nations. Less-developed nations have not been able to afford the same level of technology as their wealthier counterparts, thus creating a digital divide. While there may be some truth to this, we should not forget that market prices for technology are lower in those economies and that the growth of (mainly) Chinese exports in the 2000s and 2010s has significantly driven prices down. Open-source projects have also been instrumental in helping less-developed countries access technology at a lower cost, helping institutions save on software licenses that could otherwise become prohibitively expensive. Overall, cost has been a factor, but not the main reason the digital age's promises have not been realized.

Indeed, if cost were the only factor, one would expect that advantaged countries would have been better able to use technology to reduce inequity. My experience of living in different European Union countries has taught me otherwise. For example, when I studied in a French lycée, I was surprised to see that students were not even allowed to use technology in the classroom; the use of smartphones was strictly forbidden. In Spain, students were given a personal tablet, but this had no clear educational purpose and lacked adequate teacher training, showing how poorly planned policies can remain ineffective. This is in stark contrast to Switzerland, where technology was an integral part of the curriculum. Thoughtful integration can reduce digital divides, but coordination becomes even more critical with AI. The same can be said on a broader global scale: the role of technology and AI in education is a policy choice that varies widely across countries, districts, and institutions. Students are given different opportunities to interact with systems that are becoming more pervasive in societies, which exacerbates existing inequalities. That said, other factors such as cost, infrastructure, and cultural attitudes certainly play their part in these disparities.

Unique Challenges of AI

While some challenges that AI poses to equity in the context of education are not new, there are other aspects that differ from previous technological revolutions.

One of the biggest challenges that AI poses for education is its cost. While the Internet has operated as a (mostly) open and free platform, AI comes at a high cost that we're only beginning to understand. For example, the cost of training OpenAI's GPT-5 model is estimated at $500 million. That money is coming from investors looking for a return, and as the AI world begins

to show signs of strain, many of them will quickly start looking for ways to cash out. Are OpenAI's new ChatGPT pricing scheme (now up to $200 per month) may hint at this, but this is likely just the beginning. Ultimately, users will bear the costs.

This is a problem for everyone, and especially for low-income countries. These countries used to be able to pay at least partially for the costs of their technological infrastructure out of their own pockets because their local markets have lower costs (for example, the median price of a month of internet in the US in 2024 was about $60, while it was about $10 in Egypt). But in the world of AI, the cost of developing and running a model will be the same regardless of where the user is located. ChatGPT Pro, for example, is priced at $20 – equally, but not equitably (market-specific pricing). This is only natural in a capital-intensive market like AI, but it becomes a concern when these systems are what we base our education systems on. As costs are increasingly passed on to users regardless of their ability to pay, the digital divide can only widen.

Another challenge that is particularly pressing in the context of AI is the closed-source nature of most models. The open-source alternatives that have helped lower the cost of technology in less-developed countries are not as prevalent in the AI world. Some may point to new models like DeepSeek's V3 or R1 as examples of a new wave of open-source AI, but even these models are only open source in terms of their weights, not their training data. Researchers looking to develop more cost-effective solutions often need this data; without it, the range of possibilities is limited. For example, it is still possible to "extract knowledge" from open-source large language models (LLMs) to create smaller ones in a process known as distillation, where knowledge is transferred from a large AI model to a smaller, more efficient one. This process does not require retraining the model from scratch, but this is not always possible, and if researchers were to design an entirely new architecture, they would have to start from scratch. In general, the ability to develop cost-effective AI solutions is limited by the lack of open-source models or data, a challenge that disproportionately affects less developed countries, which have the greatest need for cost-effective solutions.

CONFRONTING BIAS IN AI SYSTEMS

Beyond cost and access, another critical challenge that threatens AI's potential for educational equity is the presence of inherent biases within LLMs, an area that I am exploring in my own research.

While Meta's Llama-3.1 might predict that "John works as a freelance" (sic), it also infers that Mary is most likely to work as a nurse and Vivek as a software engineer. These seemingly innocuous predictions can reinforce gender and racial stereotypes. Consider a student using an AI-based tutoring

system: if the system consistently suggests different career paths or learning materials based on a student's name (which, as we just saw, is a source of bias), it could subtly limit their aspirations and opportunities. These are biases that are present in the data used to train the model and are not easily corrected.

There are some techniques, such as counterfactual role reversal, that "correct" the LLM by showing it examples that challenge stereotypes. For example, it could be shown many examples of both "Mary is an engineer" and "John is a nurse" to counteract preexisting biases. However, these methods are still in their infancy and may degrade the performance of the model; after all, biases play a large role in our understanding of the world. For example, if we removed the notion of gender, we would find that they might fail to distinguish that John is (often) a "he" rather than a "she," or that breast cancer is more prevalent in females.

It might be useful to relax the definition of bias that I implicitly used above. Instead of defining "bias" based on what it is – as any difference in generated probability distributions when sensitive attributes like gender or race are changed – we can focus on what it causes and say that "bias" is any difference in generations that can be harmful to a particular group of people. This definition is more in line with the idea that bias is not inherently bad, but that it can be harmful; it is our responsibility to ensure that it is not.

The solution to this kind of bias is manifold. Public datasets exist that can be used to benchmark models for bias, and these results should be made public (model cards, similar to "nutrition facts" labels for LLMs, are a good example). This is an essential first step, although many users can be expected to ignore it – which emphasizes the need for awareness campaigns. We also need to empower researchers to develop alternatives that tackle bias at its root. One of the main hurdles is the lack of open-source training data; again, the solution is to advocate for more open-source models and data. Of course, researchers need to reciprocate by being mindful of the privacy concerns and sensitive nature of the data, such as personal information or medical records. This demands appropriate technical handling with rigorous data quality control and usage auditing.

Beyond Punitive Approaches

Fearing the potential negative consequences of AI, many educational institutions have adopted a punitive approach to its use, much like they did with the internet. I am convinced that this is not the way forward. We talked earlier about the detrimental effects of exacerbating the digital divide with a divided approach to technology, but now I would like to shift the focus to the impracticality of this approach, since much of my work has focused on the recognition of AI-generated text.

While systems designed to distinguish between human- and AI-generated text are architecturally very diverse, they all come down to the same basic idea: they look for patterns that are characteristic of AI-generated text. This approach is fundamentally flawed. The reason why is surprisingly simple: the goal of large language models is to learn the best possible approximation of the distribution of human language - in other words, to become as human-like as possible. We can think of humans as language models themselves, black boxes that spend their childhood learning the rules of language and then output text based on what they have learned, like me writing this article. If we assume that the present trend of LLMs becoming increasingly better at approximating human language continues, it is only a matter of time before they become indistinguishable from humans. Once that happens, we will not be able to rely on any system, current or future, to detect if someone is using AI.

This insight leads to an important conclusion. Educational institutions need to move beyond punitive approaches to AI use. Not only is it impractical to try to detect AI-generated text, but it can be harmful to students if they are falsely accused of using these tools. Moreover, the harm can also come from its impact on the existing digital divide, as we have seen in the previous sections. The detection of AI-generated text is therefore yet another reason to advocate for the inclusion of AI as an integral part of curricula, harnessing its potential to open up new opportunities for students around the world.

THE PATH FORWARD: BETTER AI FOR EDUCATIONAL EQUITY

If we want to promote equity in education, we need to ask for better AI. AI that is affordable, fair, and can be used by all.

To achieve this goal, I am convinced that it is essential to advocate for true open-source AI, where not only models but also training data are publicly available. It is also important to promote awareness and research on AI biases and its impact. To truly embrace AI as a tool to empower students of all backgrounds to succeed in school and in life, AI needs to be accessible across cultural and socioeconomic divides, moving away from a punitive mindset that only serves to exacerbate existing inequalities.

> ***"If we want to promote equity in education, we need to ask for better AI, one that is not only affordable but also fair and that can be used by all."***

That may sound compelling, but what exactly can Fulbrighters do? We are a diverse group of talented individuals united by the common goal of promoting cross-cultural understanding. Our global network makes us uniquely powerful to tackle AI biases and disseminate AI benefits across all countries in the program. To this end, I encourage you to consider:

Open Source Advocacy: Support or work with organizations like Apache, Hugging Face, Mozilla, or other local open source initiatives to advocate for the change you want to see. If you need a place to start, the Open Source Initiative has a list of partner organizations. Leverage the Fulbright network to coordinate advocacy efforts across multiple countries simultaneously.

Research and Awareness: Incorporate bias awareness into your teaching or research, even if it is not your primary area of expertise. Perhaps you can ask your students to compare how AI responds when they write in their native language versus English, develop educational materials on the topic, or collaborate with AI researchers globally through the program's connections. Studying bias requires cross-cultural understanding. Teams from a single country may fail to identify biases that could easily be detected by more diverse groups.

Embrace AI in Education: If you are an educator, lead the change by integrating AI literacy into your curriculum. For example, AI-powered translation tools for language learners, personalized learning platforms that adapt to individual student needs, or AI-powered chatbots that provide 24/7 student support, to name a few. Share the most and least effective practices in your experience with the broader community.

Of course, this is a long journey that requires the collaboration of educators, policymakers, students, and researchers. I will continue to work toward these goals, and I hope you can join me and other Fulbrighters in this important endeavor. Building a world where AI is a force for good in education is only possible if we work together to make it happen.

Further Reading

1. Kamalov, F., Calonge, D. S., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, 15(15), Article 11452. https://doi.org/10.3390/su151612451
2. Naous, T., Ryan, M. J., & Xu, W. (2023). Having beer after prayer? Measuring cultural bias in large language models. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 1086–1099. https://doi.org/10.18653/v1/2023.findings-acl.67
3. Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can AI-generated text be reliably detected? *arXiv*. https://doi.org/10.48550/arXiv.2303.11156
4. Solatorio, A. V., Vicente, G. S., Krambeck, H., & Dupriez, O. (2024). Double jeopardy and climate impact in the use of large language models: Socio-economic disparities and reduced utility for non-English speakers. *arXiv*. https://doi.org/10.48550/arXiv.2410.10665

Fulbright Spain grantees from a variety of programs celebrate at the 2025 pre-departure orientation in Madrid. Photo by the Fulbright Commission in Spain.

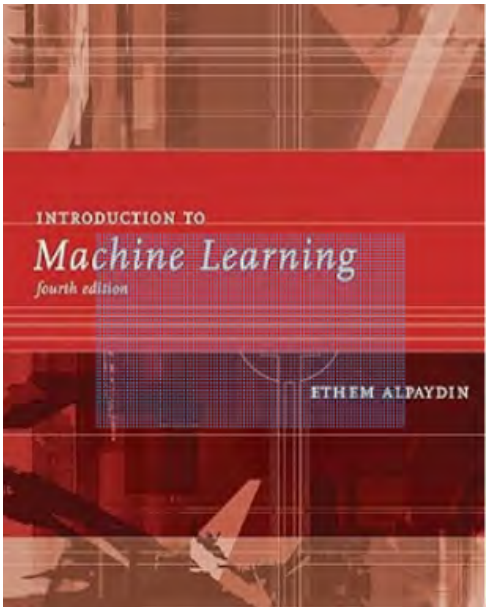
BIOGRAPHY

Aldan Creo is a Fulbright Foreign Student from Spain who will start his MS studies in Computer Science in Fall 2025. He is interested in the development of safe and fair AI systems that can benefit all, with a focus on natural language processing. His research interests include the detection of AI-generated text, hallucinations and generation artifacts; assessment of conversational risk; and multilinguality in AI. Aside from his academic pursuits, Aldan enjoys debating, traveling and volunteering, he has participated in several associations. Aldan can be reached at hello@acmc.fyi and a complete profile can be found at https://acmc.fyi/intro.

REVIEWS

The Ideal Entry Point

by Khadija Ouajjani



Introduction to Machine Learning (Revised 4th Edition), by Ethan Alpaydin who was a Visiting Fulbright Scholar to the Computer Science Institute at the University of California, Berkeley, 1997-1998.

I first stumbled upon Alpaydin's Machine Learning while compiling a concise yet solid reading list for managers and project leaders, colleagues whose work is not centered on machine learning, but whose research would unquestionably benefit from its capabilities. As an interdisciplinary researcher implementing machine learning tools within the aerospace domain, I often find that my first bottleneck isn't technical, but conversational. Many collaborators, especially those from mono-disciplinary backgrounds, engage with machine learning only through fragmented headlines or overhyped media narratives. Yet, as Research and Development offices seek to better leverage their data repositories and enhance their toolsets—often through fields outside their core expertise—the need for informed, grounded dialogue becomes urgent. In an environment saturated by polarized perspectives and oversimplified soundbites, it's rare to find an introduction to machine learning that is both technically sound and genuinely accessible, without veering into either a pedantic tone that loses the audience or a sensationalist one that gets colorful nuggets of information rather than

> ***"It's rare to find an introduction to machine learning that is both technically sound and genuinely accessible."***

grounded knowledge.

This revised and updated edition answers that need. Ethem Alpaydin, a scholar embedded in the field and capable of communicating complex ideas in layman terms, writes with precision but without condescension. He conveys foundational knowledge of machine learning clearly, often with intuitive metaphors and real-life examples. His measured tone and poised lens make the book stand out as a cogent statement of the factual capabilities of the field and its state.

One of the particular strengths of Machine Learning is the causal narrative Alpaydin uses to trace the field's evolution, linking scientific progress, data complexity, and the necessity for automated pattern discovery. Rather than presenting machine learning as a sudden revolution, he frames it as an inevitable extension of scientific inquiry itself. This framing not only helps readers grasp why machine learning emerged but also how it functions within the broader pursuit of knowledge and advancing the science behind science itself.

Alpaydin explains in logical, layman's terms, "We are now at a point where this type of data analysis can no longer be done manually, because people who can do such analysis are rare; furthermore, the amount of data is huge and manual analysis is not possible. There is thus a growing interest in computer programs that can analyze data and extract information automatically from them—in other words, learn" (37).

The chapters introducing the machine learning model types are clearly explained without complex mathematics, all the while weaving in relatable examples and rooting us in everyday challenges that would be an onerous task for humans. The well-placed mentions of artificial intelligence help in differentiating what machine learning is and isn't, and how the two intersect.

Finally, the addition of a chapter that explores the challenges and uncovered risks since the first edition (2016) is a shrewd move: It addresses data privacy; bias in data collection and manipulation; and model interpretation. This chapter also answers ethical and social questions with refreshingly succinct clarity. These have been particularly pressing in the machine learning/AI field, and Alpaydin acknowledges this has been a topic with every new technology. Echoing the sentiment of famed author Isaac Asimov, who warned that "science gathers knowledge faster than society gathers wisdom," Alpaydin considers the perils of divorcing innovation from responsibility.

Overall, Machine Learning is an ideal entry point for anyone unfamiliar with the field and needing to understand its mechanisms, capabilities, and consequences. It's concise, yet still relays the foundations of machine learning; it's thorough without burdening the pages with mathematics and algorithms, and it's grounded without downplaying the risks and pitfalls of a promising but severely unregulated field, one which can easily mine the internet and social media for data and usher in a new era of the internet.

As Alpaydin tells us, "We are already exposed to more data than what our sensors can cope with or our brains can process" (162) and thus Machine Learning will help us make sense of an increasingly complex world. If you are looking for a methodical, insightful overview of what machine learning is, its capabilities and its future, this book is the best entry point.

Ethem Alpaydin, *Machine Learning (Revised and Updated 4th Edition),* Cambridge: MIT Press, 2021. 222 pages. $15.95.

Biography

Khadija Ouajjani was a Fulbright Scholar in aerospace engineering at Wichita State University from 2016-2017. She is an intersectional researcher in the aircraft industry, currently specializing in the integration of machine learning into crashworthiness and advanced materials science. With a PhD focused on leveraging ML for defect prediction in support of aircraft certification standards, she works to bridge disciplinary gaps, collaborate on survivability-centered and numerical simulation crashworthiness projects, as well as help non-specialists engage meaningfully with machine learning tools. She can be reached at KhadijaOuajjaniPhD@proton.me

Exposing Biases in AI

by Mariofanna Milanova

Unmasking AI: My Mission to Protect What is Human in a World of Machines, by Joy Buolamwini who was a Fulbright Scholar in 2013 to Zambia.

As a Professor of Computer Science and Fulbright Scholar, I found Dr. Joy Buolamwini's book *Unmasking AI: My Mission to Protect What Is Human in a World of Machines* both inspiring and thought-provoking. The book blends personal memoir with a powerful call to action. Dr. Buolamwini shares her journey—as an African-American girl growing up in Mississippi and Tennessee, her early passion for technology and art, her experiences both as a Rhodes Scholar and a Fulbright Scholar, and finally her pioneering research at the Massachusetts Institute of Technology. It was at MIT that she exposed significant racial and gender biases in AI systems. A defining moment in her work was the discovery of the "coded gaze," when facial recognition systems failed to detect her face unless she wore a white mask, revealing the deep-rooted biases embedded in these technologies. This experiment became the foundation for the 2020 documentary *Coded Bias*, directed by Shalini Kantayya (available on Netflix). The film features Dr. Buolamwini's personal story and investigates how facial recognition systems frequently misidentify darker-skinned and female faces, raising critical concerns about algorithmic bias and its threat to civil liberties. *Coded Bias* underscores the urgent need for ethical and inclusive AI development.

Through her personal perspective, Dr. Buolamwini examines the broader impact of AI, demonstrating how these technologies can reinforce and amplify existing social inequalities. She introduces the term "excoded" to describe individuals who are marginalized or excluded as a result of biases embedded within technological systems.

"Buolamwini examines the broader impact of AI, demonstrating how these technologies can reinforce and amplify existing social inequalities."

Although this book first came out in 2023, the concept of *Coded Bias* remains highly relevant today, because, despite significant advancements in AI technology, modern AI systems continue to exhibit notable biases. Research has demonstrated that generative AI models—such as Midjourney, Stable Diffusion, and DALL·E 2—often reinforce gender and racial stereotypes, frequently portraying women and African Americans in stereotypical roles or underrepresenting them altogether. These biases contribute to the perpetuation of harmful narratives and can negatively shape public perception.

AI technologies used in critical sectors like healthcare, criminal justice, and employment also continue to mirror and amplify existing societal inequalities. Facial recognition systems, for example, have been shown to produce higher error rates when analyzing darker-skinned individuals, particularly women, resulting in misidentifications and civil rights concerns.

Buolamwini adopts an intersectional lens to examine how multiple forms of discrimination—including racism, sexism, colorism, and ableism—intersect and are embedded within AI systems. She argues that these biases are not just technical shortcomings but reflect the perspectives and limitations of the people who design and implement these technologies. While there is growing awareness and an increasing number of resources addressing bias in AI, significant challenges remain. According to Buolamwini, addressing these biases is critical to building AI systems that are fair, equitable, and inclusive—ensuring that technology serves the needs of all members of society.

The author also outlines several key strategies for addressing algorithmic bias: conducting independent algorithmic audits, developing inclusive and representative datasets, establishing comprehensive AI regulations, and fostering a culture of continuous learning. These recommendations highlight the importance of a holistic approach to AI development—one rooted in ethics, inclusivity, and accountability. By implementing such practices, we can work toward creating AI systems that serve all of humanity fairly and equitably.

I was especially inspired by the final chapter, "Cups of Hope." In this reflective section, Dr. Buolamwini emphasizes the importance of hope and perseverance, even in the face of overwhelming systemic challenges. The metaphor of "cups of hope" symbolizes small yet meaningful moments of optimism and collective action that sustain the ongoing fight for justice. This chapter stands in contrast to the harsh realities of algorithmic bias, offering

a vision for transformative change. It encourages readers to stay committed to ethical AI advocacy, despite inevitable setbacks, and to find solidarity with others striving for a more just and inclusive technological future. Through "Cups of Hope," Dr. Buolamwini reminds us that while the fight for algorithmic justice is difficult, it is not hopeless. Collective action, persistence, and shared optimism are essential to building AI systems that truly serve the needs of all people.

Towards that end, *Unmasking AI* is also a powerful call to action. Dr. Buolamwini founded the Algorithmic Justice League to address the harms caused by AI systems and to advocate for greater transparency and accountability. She invites readers from all backgrounds—not just those with technical expertise—to join the conversation about AI's role in society and to push for policies that ensure these technologies benefit everyone, not just a privileged few.

Joy Boulamwini, *Unmasking AI: My Mission to Protect What is Human in a World of Machines*. New York: Random House, 2023. 336 pages. $17.00.

Biography

Mariofanna Milanova has been a professor in the Department of Computer Science at University of Arkansas Little Rock since 2001. She received a M.Sc. in Expert Systems and Artificial Intelligence and a Ph.D. in Engineering and Computer Science from the Technical University, Sofia, Bulgaria. She also conducted post-doctoral research in visual perception in Germany. She is an IEEE Senior Member, Fulbright Scholar, and NVIDIA Deep Learning Institute University Ambassador. Her work is supported by NSF, NIH, DARPA, DoD, Homeland Security, NATO, Nokia Bell Lab, NJ, USA and NOKIA, Finland. She has published more than 120 articles, over 53 journal papers, 35 book chapters, and numerous conference papers. She also has two patents. She can be reached at mgmilanova@ualr.edu.