

The Solutions Review:  
**Navigating the  
AI Revolution**

---

Volume III | 2025







The Solutions Review:

# Navigating the AI Revolution

---

Volume III | 2025

## Imprint

---

### **Publisher**

Konrad-Adenauer-Stiftung e.V., 2025, Philippines Office  
5/F Cambridge Center Bldg., 108 Tordesillas cor. Gallardo Sts.,  
Salcedo Village, Makati City 1227 Philippines

 [info.manila@kas.de](mailto:info.manila@kas.de)

 [www.kas.de/philippinen](http://www.kas.de/philippinen)

### **Managing Editor**

Sophiya Navarro

### **Technical Editor**

Frank Cally Tabuco

### **Editorial Board**

Daniela Braun

Joshua Uel Abad

Marie Antoinette de Jesus

### **Design and Typesetting**

Lester Janz Dela Cruz

### **The Solutions Review: Navigating the AI Revolution (Volume III | 2025)**

Published annually by the Konrad-Adenauer-Stiftung e.V., Philippines Office

© 2025 Konrad-Adenauer-Stiftung e.V. All rights reserved.

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means without written permission of the Publisher, except in certain cases of fair use and noncommercial use permitted under Philippine laws.

### **Image Credits**

All image rights are retained by their respective owners. The Publisher has obtained authorization to include these images in this publication.

Cover by Lester Janz Dela Cruz with materials from Unsplash.

### **ISSN**

3116-2738 (online)

3116-272X (print)

### **Disclaimer**

The views and opinions expressed by the authors are their own and may be subject to revision. Further, these views and opinions do not necessarily represent or imply the views of Konrad-Adenauer-Stiftung (KAS) and the organizations the authors are affiliated with. Lastly, authors are solely responsible for ensuring the accuracy of facts cited in their articles.



# Contents

---

<b>Foreword</b> Daniela Braun	<b>8</b>
<b>Hybrid Intelligence: Rethinking Human-AI Systems</b> Joshua Uel Abad	<b>10</b>
<b>The AI Ecosystem in the Philippines: Growth, Challenges, and Opportunities</b> Nathan Samiley	<b>19</b>
<b>Reimagining the Future of Experiences: Exploring the Realities of AI Integration in the Philippine Tourism and Hospitality Industry</b> Eylla Laire M. Gutierrez and Ray Justin A. Villanueva	<b>24</b>
<b>Artificial Intelligence in Campaigns and Elections: A Force for Empowerment or a Threat to Democracy?</b> Marie Antoinette P. de Jesus	<b>33</b>
<b>Building AI Tools for Investigative Journalism and Public Accountability</b> Jaemark Tordecilla	<b>42</b>
<b>The Case for Sovereign AI: A Philippine Perspective</b> Carljoe Javier and Dominic Vincent Ligtot	<b>49</b>
<b>Author Profiles</b>	<b>54</b>
<b>The Editors</b>	<b>58</b>







# NAVIGATING THE AI REVOLUTION

## The Solutions Conference

The **Solutions Conference** is an annual event organized by KAS Philippines that provides a platform for experts, scholars, key stakeholders, and the public to engage in meaningful discussions of possible solutions to democracy's most pressing issues.

With artificial intelligence (AI) rapidly transforming the way we live and work, the need for cross-sectoral collaboration and timely conversations to understand and respond to its far-reaching impacts, opportunities, and risks has never been greater. To this end, KAS Philippines hosted **The Solutions Conference 2024: Navigating the AI Revolution** in Pasay City on 29-30 October 2024.

The Conference opened with plenary lectures from Mr. Dominic Ligot of CirroLytix and Data Ethics PH and Mr. Froland Tajale of the IT and Business Process Association of the Philippines (IBPAP). This was followed by a panel discussion moderated by Ms. Jacque Manabat, an independent digital journalist.

Conference day 1 also featured four dynamic parallel breakout sessions:

1. **AI, Media, and Journalism** – facilitated by Mr. Jaemark Tordecilla, Nieman Fellow for Journalism at Harvard University.
2. **Sustainable AI Workforce Development** – facilitated by Mr. Sonny Carlos of the Analytics and AI Association of the Philippines (AAP).
3. **AI for All: Bridging the Digital Divide for a Better Future** – facilitated by Ms. Gina Romero of Connected Women.
4. **Harnessing AI for Good Governance** – facilitated by Ms. Lei Motilla and Mr. Julius Sambo of AI4GOV.

A panel discussion featuring the speakers from each breakout session concluded the first day of the activity.

On day 2 of the Conference, Mr. Dominic Ligot led a **training session on prompt engineering**, designed to equip participants with the skills to craft effective prompts and maximize the use of AI models, particularly generative AI, in their daily lives.

The two-day conference brought together industry leaders, tech experts, innovators, and the public to learn more about AI and explore how this technology can be a powerful catalyst for social good and drive innovative solutions to critical global challenges.

# The Publication

**The Solutions Review** presents a carefully curated compendium of think pieces that bring together the key themes and different perspectives shared by experts, scholars, and key stakeholders during the annual Solutions Conference.

The **2025 issue** will explore the theme **“Navigating the AI Revolution”** and build upon the discussions from The Solutions Conference 2024 hosted by KAS Philippines on this very topic.

KAS Philippines hopes that The Solutions Review will serve as invaluable reference material for a broader audience of passionate individuals and groups, enabling them to further drive positive and meaningful progress on the thematic issue discussed during the Conference.

Read past issues of The Solutions Review at <https://www.kas.de/en/web/philippinen/publications>.



# Foreword

**“We want AI to be a force for good. We want an AI where everyone collaborates and everyone benefits.” — Ursula von der Leyen, European Commission President, during the AI Action Summit in Paris, February 2025.**

Artificial intelligence or AI is no longer a concept of the future—it is the defining force of the present. It is not only reshaping the ways we work and communicate but also revolutionizing healthcare, education, journalism, defense, and governance. AI has rapidly moved from laboratories all the way to the center of public life—and this is precisely the reason why KAS, as a political organization, is working on AI and its broader implications to society. This moment represents not merely a technological shift but a societal transformation. And as with any revolution, it comes with opportunities, risks, and profound responsibilities.

AI has the ability to significantly reshape politics—changing how governments interact with citizens, make decisions, and allocate public resources. AI-powered tools can enhance policymaking and accountability through real-time data analysis, predictive modelling, and more responsive public services. Campaigns can become more targeted and personalized, while AI-driven platforms may increase civic participation through automated information delivery and feedback systems. At the same time, these technologies raise urgent concerns about algorithmic bias, political manipulation, and the concentration of power—making thoughtful regulation and democratic oversight more critical than ever.

The rapid deployment of generative AI systems, autonomous decision-making tools, and algorithmic infrastructure in public and private sectors has sparked debate. While some view AI as a lever for productivity, innovation, and inclusion, others raise concerns about surveillance, algorithmic bias, misinformation, and job displacement.

*The Solutions Review 2025: Navigating the AI Revolution* is a follow-up publication of the KAS Solutions Conference held in October 2024 in Metro Manila. While the Conference and this publication approach the



*Daniela Braun (left) with Jacque Manabat (right) during the 2024 KAS Solutions Conference.*

future of AI with solutions and optimism, this volume does not seek to either idealize or vilify it. Instead, this work explores the multifaceted challenge of ensuring that AI serves the public good. It addresses key policy questions: How do we align AI development with democratic values? What frameworks are needed to promote transparency and accountability? How can we foster innovation while safeguarding rights and liberties?

This publication brings together a diverse set of voices, including political scientists, journalists, economists, and technologists, to examine the far-reaching implications of AI and to offer grounded, forward-thinking policy solutions. What emerges from their contributions is not a single narrative, but a dialogue that reflects the complexity and urgency of the AI revolution.

As the AI landscape continues to evolve with unprecedented speed, *The Solutions Review 2025* aims to serve as a resource for policymakers, experts, civil society actors, and interested citizens—offering not only analysis, but also a call to action. The AI revolution is here. Our role now is to shape it wisely with clarity of purpose, a steadfast commitment to safeguarding democratic values, and an unwavering respect for human dignity.

We at KAS Philippines are confident that the ideas and solutions presented in the following pages can contribute to navigating the AI revolution. Let us all work together for a sustainable, secure, innovative, democratic, and prosperous future.

**Daniela Braun**  
Country Director  
KAS Philippines



*Joshua Uel Abad during the 2024 KAS Solutions Conference.*

# Hybrid Intelligence: Rethinking Human-AI Systems

**Joshua Uel Abad**

---

Imagine a high school teacher preparing for a social science class. Traditionally, this teacher would spend hours crafting lesson plans, researching supplementary materials, and anticipating student questions. Today, with publicly accessible generative AI tools like ChatGPT, there is growing potential to leverage AI to assist in generating customized lesson outlines, drafting engaging content, and simulating learner interactions. However, realizing this potential is not automatic. Publicly available AI models are typically designed for general use and will require significant customization, training, and pedagogical guidance to be truly effective in the classroom. In the absence of such modifications, the outputs may lack accuracy, contextual sensitivity, or alignment with curriculum standards. At the same time, the meaningful integration of AI in education also requires institutional support and change management, ensuring that tools complement rather than complicate a teacher's role.

A similar picture emerges in healthcare. AI models trained to detect anomalies in medical images can support radiologists by filtering vast volumes of data and flagging potential concerns. Yet, these systems must be trained on high-quality, domain-specific datasets and built with expert oversight. Even then, their contributions remain limited without the radiologist's clinical expertise, contextual understanding, and ethical judgment. Integrating AI into diagnostic practice is as much about organizational readiness and workflow reconfiguration as it is about model accuracy.

The examples above reflect a broader truth: while AI can alleviate routine or repetitive tasks, its ability to contribute meaningfully to complex decision-making depends on intentional system design, expert training,

and human-centered integration. Such cases also illustrate the concept of hybrid intelligence—the idea that humans and AI systems can collaborate to achieve outcomes greater than either could alone. Nevertheless, this framing invites a deeper question: what does it mean to “collaborate” with AI? In common usage, collaboration implies mutual awareness, shared goals, and interactive agency. However, AI systems lack consciousness and sentience. Thus, calling AI a collaborator may risk anthropomorphizing the technology or masking underlying power dynamics in human-machine systems. To move beyond metaphor, we must unpack the different modalities of collaboration emerging in practice.

### Defining human-AI collaboration

One useful framework for understanding these modalities of human-AI collaboration includes three conceptual layers: embedded actor, co-agency, and editorial augmentation.

As an embedded actor, AI functions as a tightly integrated system component with delegated authority over specific, bounded tasks such as anomaly detection or recommendation sorting. In this role, AI does not “collaborate” in a traditional sense but operates within human-designed decision environments, shaped by rules, goals, and oversight.<sup>1</sup>

As co-agents, humans and AI systems share control over a task, with AI contributing distinct strengths such as pattern recognition, contextual reasoning, or strategic foresight, in a process cognitive scientists refer to as distributed cognition.<sup>2</sup> This framework is particularly relevant in fields like finance or logistics, where real-time data analysis and human judgment must converge to drive action.

Finally, in editorial augmentation, AI systems generate raw content, suggestions, or variations, with human users acting as curators in charge of refining, contextualizing, and ethically guiding outputs toward meaningful outcomes.<sup>3</sup>

This nuanced understanding of human-AI collaboration has two important implications. First, it calls for more precise language in how we describe AI and its role, not as a peer or an autonomous partner, but as a catalytic system that

must be designed with transparency, alignment, and co-adaptability. Second, it points to the need for hybrid systems that support human strengths, including empathy, creativity, and moral reasoning, while compensating for cognitive limitations. When collaboration is thoughtfully engineered through complementary design, transparent governance, and shared accountability, hybrid intelligence becomes more than a technological upgrade, but a new way of doing things.

### What works and what does not?

Research from the MIT Sloan Management Review highlights that human-AI teams excel in tasks involving content generation, pattern recognition, and large-scale data analysis.<sup>4</sup>

In journalism, AI can generate initial drafts, analyze datasets, and suggest headlines while human journalists provide context, fact-checking, and narrative coherence.<sup>5</sup>

In healthcare, AI algorithms analyze medical images or patient data to identify potential diagnoses or treatment options. However, human clinicians interpret these findings within the broader clinical context, considering the patient's history, preferences, and ethical considerations. This collaboration enhances diagnostic accuracy and patient outcomes.<sup>6</sup>

Similarly, in the financial services sector, AI models are used to detect fraud, flag unusual transactions, and identify market anomalies. Human analysts then interpret these signals, assess contextual risks, and make strategic decisions based on broader economic and organizational insights. This division of labor reflects a growing trend toward hybrid intelligence in finance, where machine efficiency is paired with human expertise to improve decision-making quality.<sup>7</sup>

Within these different contexts, the division of labor leverages AI's ability to process vast amounts of information quickly while relying on humans to provide a nuanced understanding of the audience and navigate ethical complexities, resulting in improved efficiency and outcomes. These examples illustrate how hybrid intelligence can harness complementary capabilities when designed with care and applied to tasks well suited to AI augmentation.

However, effectiveness is not guaranteed across all domains. Research published by the MIT Sloan Management Review warns that in contexts involving ambiguous decision-making or requiring deep contextual and ethical understanding, human-AI teams do not consistently outperform human judgment alone.<sup>8</sup> For example, in areas such as legal sentencing or hiring, AI systems trained on biased data may perpetuate discrimination, especially when their recommendations are applied uncritically or without appropriate oversight.<sup>9</sup> These limitations do not invalidate the potential of hybrid intelligence but rather underscore the importance of thoughtful design, domain-specific training, and ongoing human judgment in high-stakes applications involving AI.

In this regard, the Wharton School emphasizes that hybrid intelligence is not a one-size-fits-all solution.<sup>10</sup> Success depends on what is described as “task-fit” thinking or the strategic division of cognitive labor between humans and AI. Tasks that are well-defined, repetitive, and data-intensive are often well-suited for AI assistance or automation. In contrast, tasks that rely on ethical judgment, empathy, or open-ended creativity tend to require stronger human involvement.<sup>11</sup> However, this distinction is not always clear-cut.

In practice, many tasks contain overlapping elements such as data-driven components interwoven with ethical or interpretive dimensions. For example, content moderation, medical triage, or hiring decisions all involve large-scale data processing but also require values-based reasoning and contextual sensitivity. These “blended” tasks are particularly susceptible to problems like biased training data, which can amplify systemic inequalities, particularly if AI outputs are applied without careful human oversight. As such, “task-fit” thinking should not be seen as a rigid classification, but rather as a dynamic and iterative process—one that considers not only the nature of the task but also the risks, consequences, and social context surrounding its implementation. A more holistic approach to hybrid intelligence requires recognizing and designing for these interdependencies, ensuring that both AI capabilities and human judgment are engaged where they matter most.

Within this context, AI researcher Jessy Lin advances the concept of the “editorial mindset” as a

new form of knowledge work within hybrid systems. In creative domains, AI can generate ideas, drafts, or prototypes at scale, offering new possibilities for content creation. However, these outputs still require human curation, editing, and contextual alignment to meet strategic, cultural, and ethical goals.<sup>12</sup> For example, in advertising, AI might suggest multiple campaign concepts. Still, human marketers can refine those ideas and come up with marketing campaigns that resonate with brand values, audience expectations, and social sensitivities.

This “editorial” approach preserves human agency and quality control while leveraging the efficiency and generative capacity of AI. However, it also raises important ethical concerns. Specifically, many AI systems are trained on vast datasets that include human-created works, some of which may be subject to copyright protection, data privacy laws, or cultural ownership. The use of such materials without consent or attribution raises unresolved questions about intellectual property and fairness. Moreover, as generative AI becomes more prevalent, there is a growing risk of labor displacement—particularly for artists, designers, and writers whose contributions have enabled the development of these tools but who may now find their roles undermined by AI.

A critically informed “editorial” mindset, therefore, must go beyond content refinement. It must also include awareness of the origins of training data, sensitivity to power dynamics in the creative and labor economy, and a commitment to ethical sourcing and attribution. By embedding these considerations into hybrid creative processes, we ensure that the collaboration between humans and AI is not only productive but also just and accountable. The lesson is that human-AI collaboration works best when systems are designed to respect and leverage the unique strengths and limitations of each. Understanding what works and what does not is essential to advancing hybrid intelligence that is both effective and ethical.

### **Augmentation vs. automation**

The early narrative around artificial intelligence centered on automation—machines replacing humans in routine, repetitive tasks to increase efficiency and reduce costs. While automation

remains an important application in some cases, the emerging paradigm increasingly emphasizes augmentation or using AI to enhance human capabilities and creativity rather than displacing them.<sup>13</sup> Augmentation reframes how we think about productivity and value. Instead of measuring success solely by output or speed, augmentation considers how AI empowers humans to engage in higher-order cognitive, strategic, and creative work. This shift aligns with the idea of hybrid intelligence, in which humans and AI systems form teams with complementary strengths.

However, augmentation is not without its limitations and challenges, particularly in how humans interact with AI systems. In many cases, AI models function as “black boxes”—their internal logic and decision-making processes are opaque, even to their developers. This lack of explainability can lead to low trust in AI outputs, especially in high-stakes domains like healthcare. Studies have shown that clinicians often hesitate to adopt AI systems whose outputs they cannot interpret or challenge, even if those systems perform well statistically.<sup>14</sup> The lack of transparency in AI decision-making thus becomes a major barrier to integration and trust.

Moreover, poor human-AI interaction not only hinders trust. It can also lead to automation bias, where users defer too readily to AI recommendations, even in the face of contradictory evidence.<sup>15</sup> Conversely, algorithm aversion occurs when humans ignore or underutilize AI tools due to a lack of understanding or past negative experiences.<sup>16</sup> Effective augmentation requires calibrated trust, and users must be trained to engage critically with AI, understanding both its strengths and its limitations.

Still, these concerns raise further questions that merit deeper exploration: How do we ensure that AI remains a tool for empowerment, rather than a crutch for dependency? As AI systems become more capable, do we risk over-reliance, where human users unlearn essential skills or defer decision-making to machines by default? Research in education and aviation has shown that overdependence on intelligent systems can result in skill degradation and lower situational awareness.<sup>17</sup> Designing for augmentation thus requires more than just algorithmic performance—it demands robust human-machine interfaces,

effective user training, and organizational cultures that promote human agency within AI systems. In sectors like healthcare, law, and education, where ethical reasoning, empathy, and discretion are essential, AI should assist, not replace, human skills and input. Nevertheless, this reframing invites reflection. While embracing AI as a collaborator opens new pathways for development and progress, it also demands intentionality, ethical design, and critical awareness at every step.



### Rethinking organizational systems

Successfully implementing hybrid intelligence requires organizations to move beyond technological adoption toward structural, cultural, and strategic transformation. As AI becomes more deeply embedded into daily operations, organizations must evaluate their readiness for human-AI collaboration, which begins with a clear-eyed assessment of current workflows, capabilities, and objectives. Organizational readiness involves understanding the current state of operations: Where are decisions delayed due to manual processing? Which workflows are constrained by repetitive tasks or information bottlenecks? For many firms, especially in knowledge-intensive sectors, these inefficiencies may be tied to legacy systems, fragmented data management, and siloed teams. AI has the potential to address such problems by automating routine processes, surfacing hidden patterns in data, and enabling more agile decision-making.<sup>18</sup>

However, these benefits do not emerge automatically. They require deliberate process redesign grounded in “complementarity by design” or structuring tasks in a way that humans can interpret, challenge, and override AI outputs to ensure accountability and contextual awareness.<sup>19</sup> Equally important is the need to set clear, measurable objectives for human-AI integration.

This includes defining metrics of success such as improved decision speed, reduction in error rates, cost efficiency, or gains in service quality. Organizations should identify specific pain points where AI could create value and pair these with target benchmarks to assess performance over time. Without such clarity, AI adoption risks becoming a technology-first exercise rather than a human-centered transformation.

As such, a critical part of redesigning organizational systems is engaging the people directly affected by these changes. Employee feedback is indispensable, not only for technical calibration but for shaping trust, buy-in, and long-term sustainability. Workers often possess deep tacit knowledge of where systems break down, what tasks are meaningful, and where automation may enhance or hinder performance. Involving them early in the implementation process can uncover unintended consequences and foster co-ownership of AI-driven change.<sup>20</sup>

Beyond system design and metrics, organizational culture must also evolve to support hybrid intelligence. Trust in AI is essential, but it must be calibrated, not blind, noting that both overreliance and total rejection of technologies can lead to failure.<sup>21</sup> To this end, ensuring trust involves building AI systems that are explainable and transparent, allowing users to understand how decisions are made and when they should be questioned.<sup>22</sup> These systems should include feedback loops that continuously learn from human input, improving accuracy and contextual relevance over time. Cultivating a culture of psychological safety, where employees feel empowered to question AI outputs, raise concerns, and suggest improvements, is equally crucial. Training programs that build AI fluency among employees can bridge the gap between technical capabilities and everyday use, enabling more meaningful human-AI interaction and fostering a shared language around performance, limitations, and expectations. Hybrid systems must also be continuously evaluated for bias, privacy risks, and skill erosion. When humans are removed from critical processes or become overly reliant on AI, there is a risk of deskilling and diminished human judgment over time. Ethical governance and human-centered design are therefore essential not only at the point of implementation but throughout the AI system's lifecycle.<sup>23</sup>



Photo taken during a lecture at the KAS Solutions Conference 2024.

## Towards equitable collaboration

By rethinking organizational systems through a participatory, goal-oriented, and ethically grounded lens, AI can be integrated not as a disruptive force but as a collaborative mechanism, one that elevates both institutional performance and human agency. To ensure that human-AI collaboration delivers broad-based benefits rather than exacerbating inequalities or harms, policymakers must move beyond reactive regulation toward proactive, design-oriented governance. This approach requires comprehensive policies that foster AI literacy, mandate meaningful human oversight, incentivize hybrid workflows, support workforce transitions, and establish robust ethical auditing frameworks. Together, these measures can create an ecosystem where AI augments human potential while safeguarding social values and equity.

### AI literacy initiatives

AI literacy initiatives are foundational to equitable collaboration. Although AI systems have become increasingly ubiquitous in workplaces and daily life, public understanding of AI's capabilities and limitations remains uneven, heightening the risks of misuse, mistrust, and a widening digital divide. Comprehensive public education campaigns are essential to demystify AI, emphasizing its role as a collaborator rather than a tool to replace humans. Embedding AI literacy in school curricula equips future generations with critical skills to navigate AI-augmented environments effectively. Similarly, corporate training programs must include AI literacy to prepare workers for hybrid work environments.

Recent research highlights that AI literacy enhances human-AI team performance by enabling users to

calibrate trust appropriately and engage critically with AI outputs.<sup>24</sup> This is particularly evident in settings where users must evaluate AI-generated suggestions within specific ethical, cultural, or procedural contexts. For instance, using the previous cases mentioned in this paper, teachers using AI-powered learning platforms are better able to tailor instruction when they understand how recommendations are generated. Similarly, in healthcare, health professionals with training in AI are more effective at integrating AI-based diagnostic support while maintaining clinical judgment. In administrative and public service settings, AI-literate staff are more likely to spot anomalies, interpret automated assessments with caution, and escalate concerns when outputs conflict with situational realities. Across these domains, AI literacy fosters a balanced relationship between trust and oversight, reducing the risk of both automation bias and human disengagement.

Additionally, AI literacy initiatives must be inclusive and culturally sensitive, addressing disparities in access and educational resources. Governments and institutions should collaborate with community organizations to tailor programs for diverse populations, ensuring that marginalized groups are not left behind in the AI transition. International cooperation can facilitate knowledge sharing and capacity building, particularly for developing economies that are facing rapid AI adoption without sufficient infrastructure.

As highlighted in the Human-AI Collaboration Framework developed by the nonprofit organization Partnership on AI in 2019, transparency and user understanding are crucial to building trust and promoting responsible AI adoption across various contexts.<sup>25</sup> Thus, AI literacy is not merely technical training but a societal imperative for informed participation in the AI-augmented future.

### **Human-in-the-loop mandates**

Human-in-the-loop mandates are critical for preserving accountability, ethical integrity, and contextual judgment in AI applications, especially in high-stakes domains such as healthcare, criminal justice, and legal decision-making. These mandates require that AI systems operate under meaningful human oversight, ensuring that final decisions remain with individuals who can interpret, question, and override algorithmic outputs. This approach is

designed to mitigate the risks of bias, technical error, or opaque reasoning that could otherwise result in harm. Although human-in-the-loop frameworks are often presented as safeguards, human oversight does not automatically eliminate bias. Humans are also susceptible to cognitive shortcuts, overreliance, or dismissiveness toward machine recommendations. As such, what matters most is the quality of the oversight: Are professionals trained to interrogate algorithmic decisions? Do they understand how the AI model was trained and what limitations it carries? Without proper AI literacy and institutional support, human oversight can be symbolic rather than substantive.

The ethical considerations surrounding these mandates also extend to how data is collected, processed, and used to train AI systems. Health data, for example, is often deeply personal, longitudinal, and collected in highly unequal systems. Ensuring that data is sourced and used ethically means establishing robust consent mechanisms, secure data infrastructure, and transparency about how models evolve. These requirements are not just technical; they must be encoded into regulatory and organizational norms.

Human-in-the-loop mandates are not just static layers added to AI systems, but a dynamic and evolving role that must be designed in tandem with institutional workflows and ethical governance. Therefore, implementing these frameworks demands cultural and operational shifts. Organizations must foster environments where human judgment is valued alongside AI efficiency, and where questioning algorithmic outputs is not seen as resistance but as responsible practice. Calibrated trust is essential because hybrid systems fail when users either unquestioningly accept or reflexively reject AI recommendations. Policymakers and developers should prioritize explainable AI systems that provide uncertainty indicators, model rationales, and audit trails, supporting informed human decision-making.<sup>26</sup> Sector-specific guidelines are also necessary, ensuring that human-in-the-loop mandates are supported by policy frameworks that balance innovation, user safety, and ethical integrity.

### **Tax incentives**

Tax incentives can serve as a powerful policy lever to accelerate the adoption of AI systems that prioritize augmentation rather than automation.

When structured thoughtfully, these incentives can shift corporate behavior toward human-centered AI integration by encouraging investments that enhance, rather than displace, human labor. For example, tax credits could be extended to organizations that develop or deploy AI tools with built-in human-in-the-loop mechanisms, or those that demonstrate integration of AI into workflows that elevate task complexity, support decision-making, or reduce cognitive overload for workers. Similarly, companies that fund employee reskilling, cross-functional AI training, or research into collaborative AI interfaces should be rewarded for fostering the conditions necessary for hybrid intelligence to thrive. As studies show, firms that adopt augmentation-focused strategies tend to report gains in productivity, innovation, and employee satisfaction—benefits that align private sector outcomes with broader public goals.<sup>27</sup>

To ensure these incentives truly support human-AI collaboration, governments must go beyond general AI subsidies and adopt a more targeted, criteria-driven approach. Public benefits such as tax breaks, grants, or procurement eligibility should be tied to explicit indicators of augmentation. These may include evidence of ongoing workforce development programs, measurable improvements in job quality, such as employee retention, stress reduction, skill utilization, and demonstrable human oversight of AI outputs. By conditioning incentives on these factors, policymakers can prevent unintended consequences, including incentivizing job displacement under the guise of innovation. This approach also reframes AI development not merely as a technical upgrade, but as an opportunity to reimagine how work is organized and supported across industries.

Governments can further strengthen these efforts by establishing independent auditing or certification systems to verify whether publicly supported AI systems meet augmentation goals. For instance, an AI augmentation label could certify tools that promote human agency, transparency, and explainability. This label, akin to certifications for environmental or labor standards, would not only foster consumer and investor trust but also differentiate responsible innovation in a rapidly expanding market. Public procurement policies can also be aligned with these values by prioritizing vendors and

contractors whose AI solutions reflect human-centered design principles. In doing so, the state can act as a market shaper, using its buying power to reinforce standards for ethical and inclusive AI adoption.

Finally, accountability mechanisms must accompany financial incentives to ensure transparency and continuous improvement. Firms receiving tax benefits or public funds for AI projects should be required to report regularly on workforce impacts, including metrics related to job creation or displacement, training and upskilling initiatives, and the role of human oversight in AI workflows. These disclosures will allow regulators and researchers to track whether augmentation goals are being met, inform future refinements to incentive structures, and encourage organizations to reflect on their internal practices and consider the long-term societal value of their AI strategies. By shaping market incentives in this direction, governments can help steer AI development toward a future where technological progress amplifies human potential and supports more equitable and meaningful forms of work.



*Mr. Froland Tajale of the IT & Business Process Association of the Philippines (IBPAP) delivers a plenary lecture on how the Philippine IT-BPM sector is adapting to AI at the 2024 KAS Solutions Conference.*

### Workforce transition funds

Hybrid systems create new roles that require skills in managing, interpreting, and collaborating with AI systems; however, many workers currently lack these competencies. As such, transition funds are essential to support workers who are displaced or whose roles have been transformed by the integration of AI in the workplace. These dedicated funds can finance reskilling and upskilling programs, focusing on digital literacy, AI interaction skills, and domain-specific knowledge needed for hybrid workflows. Such investments can help improve employment outcomes and

service effectiveness when combined with human-AI collaboration frameworks. Transition funds also serve as safety nets during periods of labor market adjustment, thereby reducing social disruption and inequality.

Effective workforce transition policies necessitate partnerships among governments, educational institutions, and industry to align training with evolving labor market demands. Programs should be accessible, affordable, and tailored to diverse worker populations, including marginalized groups disproportionately impacted by AI automation. Continuous learning models that integrate on-the-job training with formal education can foster adaptability in rapidly changing environments. International examples show that proactive workforce development mitigates resistance to AI adoption and enhances public trust in technology-driven change.<sup>28</sup> Thus, transition funds are not only economic instruments but also social investments in inclusive AI futures.

### **Ethical auditing bodies**

Auditing bodies provide independent oversight to ensure that human-AI decision systems operate fairly, transparently, and without bias. These entities conduct systematic reviews of AI algorithms, data sets, and deployment practices, identifying ethical risks and recommending corrective actions. Given the complexity and opacity of many AI systems, external audits are crucial for maintaining public confidence and preventing harm such as discrimination or privacy violations. Recent policy proposals advocate for establishing such bodies with multidisciplinary expertise, including ethicists, technologists, and social scientists, to evaluate AI impacts holistically.<sup>29</sup> Ethical auditing complements human-in-the-loop mandates by scrutinizing the underlying AI design and data governance.

To be effective, auditing bodies must have apparent authority, transparency in their processes, and mechanisms for enforcement. They should engage with affected communities to understand real-world consequences and incorporate diverse perspectives. International coordination can harmonize standards and share best practices, thereby preventing regulatory arbitrage and fostering global coherence in AI governance.<sup>30</sup> Moreover, audits should be

ongoing rather than one-time assessments, reflecting the dynamic nature of AI systems and their evolving societal roles. By institutionalizing ethical oversight, policymakers can ensure that the development of hybrid intelligence aligns with human values and social welfare.

### **Conclusion: it is you, not AI**

Beyond technical and organizational considerations, hybrid intelligence raises profound philosophical and emotional questions about what it means to be human in an AI-augmented world. Human intelligence embodies irreplaceable qualities: moral reasoning that navigates complex ethical dilemmas, empathy that connects us to others' experiences, and creativity born from lived experience and cultural context. These dimensions cannot be fully replicated by AI.<sup>31</sup>

In hybrid systems, the human role is strategic and editorial. As academic research and practical experience converge, it becomes clear that thoughtful policy design is indispensable to harnessing AI as a mechanism that amplifies human creativity, judgment, and dignity. This reframing calls for a new humanism that values human uniqueness while embracing technological augmentation. We must ask: What kind of intelligence do we want to cultivate in ourselves and our machines? The answer shapes the future of work, society, and human flourishing.

Not all tasks benefit from AI, but many can—if collaboration is thoughtfully designed. Hybrid intelligence provides a robust framework for integrating human creativity and judgment with AI's computational capabilities. By leveraging AI, we can unlock new possibilities for innovation, equity, and meaning in both work and life. Any technology, no matter how advanced, is a tool shaped by human choices, values, and creativity. Hybrid intelligence offers immense promise, but its success depends not on the capabilities of AI alone but on how we, as individuals and societies, choose to wield it. Ultimately, humans, not AI, determine the direction, meaning, and impact of this collaboration. The challenge is not to fear AI but to become smarter, more reflective, and more humane in how we integrate it into our lives and use it as a catalyst for progress, innovation, and meaning.

## References

- 1 Lucy Suchman, *Human-Machine Reconfigurations: Plans and Situated Actions, 2<sup>nd</sup> ed.* (Cambridge University Press, 2007).
- 2 James Hollan, Edwin Hutchins, and David Kirsh, "Distributed cognition: toward a new foundation for human-computer interaction research," *ACM Transaction on Computer-Human Interaction* 7, no. 2 (2000): 174-196, <https://doi.org/10.1145/353485.353487>.
- 3 Jessy Lin, "Rethinking Human-AI Interaction," June 8, 2020, <https://jessylin.com/2020/06/08/rethinking-human-ai-interaction/>.
- 4 Jon McCormack, Toby Gifford, and Patrick Hutchings, "Autonomy, Authenticity, Authorship and Intention in computer generated art," *EvoMUSART 2019: 8th International Conference on Computational Intelligence in Music, Sound, Art and Design*, March 2019, <http://dx.doi.org/10.48550/arXiv.1903.02166>.  
MIT Sloan Office of Communications, "Humans and AI: Do they work better together or alone?," *MIT Sloan*, October 28, 2024, <https://mitsloan.mit.edu/press/humans-and-ai-do-they-work-better-together-or-alone>.
- 5 Thomas H. Davenport and Rajeev Ronanki, "Artificial Intelligence for the Real World," *Harvard Business Review*, January-February 2018, 108-116.
- 6 Eric J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine* 25, no. 1 (2019): 44-56, <https://doi.org/10.1038/s41591-018-0300-7>.
- 7 Davenport and Ronanki, "Artificial Intelligence," 2018.
- 8 "Humans and AI," *MIT Sloan*, 2024.
- 9 Ibid.
- 10 Cornelia C. Walther, "Why Hybrid Intelligence Is the Future of Human-AI Collaboration," *Knowledge at Wharton*, March 11, 2025, <https://knowledge.wharton.upenn.edu/article/why-hybrid-intelligence-is-the-future-of-human-ai-collaboration/>.
- 11 Ajay Agrawal, Joshua Gans, and Avi Goldfarb, *Prediction Machines: The Simple Economics of Artificial Intelligence*, (Harvard Business Review Press, 2019).
- 12 Lin, "Human-AI," 2020.
- 13 H. James Wilson and Paul R. Daugherty, "Collaborative Intelligence: Humans and AI Are Joining Forces," *Harvard Business Review*, July-August 2018, 114-123.
- 14 Sana Tonekaboni, Shalmali Joshi, Melissa D. McCradden, Anna Goldenberg, "What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use," *Proceedings of Machine Learning Research* 106, (2019): 359-380, <https://proceedings.mlr.press/v106/tonekaboni19a.html>.
- 15 Raja Parasuraman and Victor Riley, "Humans and Automation: Use, Misuse, Disuse, Abuse," *Human Factors*: 39, no. 2 (1997): 230-253, <https://doi.org/10.1518/001872097778543886>.
- 16 Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey, "Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err," *Journal of Experimental Psychology: General* 144, no. 1 (2015): 114-126, <https://doi.org/10.1037/xge0000033>.
- 17 Lisanne Bainbridge, "Ironies of automation," *Automatica* 19, no. 6 (1983): 775-779, [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8).  
Mica R. Endsley, "From Here to Autonomy: Lessons Learned From Human-Automation Research," *Human Factors* 59, no. 1 (2017): 5-27, <https://doi.org/10.1177/0018720816681350>.
- 18 Jacques Bughin, Eric Hazan, Sree Ramaswamy, Michael Chui, Tera Allas, Peter Dahlström, Nicolaus Henke, Monica Trench, "ARTIFICIAL INTELLIGENCE: THE NEXT DIGITAL FRONTIER?," *McKinsey Global Institute*, June 2017, <https://www.mckinsey.com/de/~media/mckinsey/industries/advanced%20electronicsour%20insights/how%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-discussion-paper.pdf>.
- 19 Thomas H. Davenport and Julia Kirby, *Only Humans Need Apply: Winners and Losers in the Age of Smart Machines*, (Harper Business, 2016).
- 20 Erik Brynjolfsson, Daniel Rock, and Chad Syverson, "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," in *The Economics of Artificial Intelligence: An Agenda*, edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, (University of Chicago Press, 2023), 23-56.
- 21 Ibid.
- 22 Finale Doshi-Velez and Been Kim, "Towards A Rigorous Science of Interpretable Machine Learning," *arXiv preprint*, (2017), <https://doi.org/10.48550/arXiv.1702.08608>.
- 23 Kate Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, (Yale University Press: 2021).
- 24 James Hutson, "Human-AI Collaboration in Writing: A Multidimensional Framework for Creative and Intellectual Authorship," *International Journal of Changes in Education*, (2025): 1-10, <https://doi.org/10.47852/bonviewIJCE52024908>.
- 25 "Human-AI Collaboration Framework & Case Studies," *Partnership on AI*, September 25, 2019, <https://partnershiponai.org/paper/human-ai-collaboration-framework-case-studies/>.
- 26 Cynthia Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence* 1, no. 5 (2019): 206-215, <https://doi.org/10.1038/s42256-019-0048-x>.
- 27 Raymond Uzwyshyn, "Human-AI Ph.D Level Co-Intelligence 2025-2035: Roadmaps, Reflections and Recommendations for Human AI Symbiosis," LinkedIn, April 27, 2025, <https://www.linkedin.com/pulse/humanai-phd-level-co-intelligence-2025-2035-roadmaps-uzwyshyn-ph-d--n1h3c/>.
- 28 "Policy Recommendation Concerning the Use of Artificial Intelligence," *International Transport Forum*, May 22, 2025, <https://www.itf-oecd.org/transport-ministers-agree-groundbreaking-policy-recommendations>.
- 29 Ibid.
- 30 Ibid.
- 31 Luciano Floridi, *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, (Oxford University Press, 2019).



*Nathan Samiley during the 2024 KAS Solutions Conference.*

# The AI Ecosystem in the Philippines: Growth, Challenges, and Opportunities

**Nathan Samiley**

---

Artificial intelligence (AI) is rapidly transforming global economies, redefining industries, and reshaping the skills required for the workforce of the future. From automating routine tasks to enabling advanced predictive analytics and decision-making, AI has become a critical driver of productivity and innovation.

In its 2025 Human Development Report, the United Nations Development Programme (UNDP) warns of the slowing momentum in global human development. Heading into 2020, global indicators suggested that we are progressing steadily toward achieving “a very high human development index (HDI) world by 2030.” However, major global setbacks, most notably the 2020 COVID-19 pandemic, disrupted this positive trajectory. Data showed that recovery between 2021 and 2024 had been slow and sluggish, threatening long-term global development goals. If this trend continues, reaching global HDI targets will be delayed, not just by a few years, but by decades. In response, the UNDP, in its report, emphasizes the importance of global AI readiness, recognizing the potential of this technology as a catalyst for progress, and highlights that countries such as the Philippines are well-positioned to build a strong momentum in leveraging AI for inclusive and sustainable development.<sup>1</sup>

Countries that can strategically adopt and govern AI technologies will reap substantial benefits in economic competitiveness and social development. However, for the Philippines, embracing AI is both a challenge and an opportunity. With a young, tech-savvy population, a vibrant services sector, and a strong diaspora footprint in technology, the country is uniquely positioned to benefit from AI adoption. Still, challenges such as digital infrastructure gaps, limited technical expertise, and regulatory uncertainty threaten the country’s progress if not adequately addressed. In this context, a coordinated, inclusive, and ethical approach to AI development is essential. The emergence of a robust and competitive Philippine AI ecosystem—composed of government, academia, and the private sector—offers a path forward.

## Key actors in the Philippine AI ecosystem

### 1 The government: enabling policy and strategic direction

The Philippine Government views AI as a national development priority. Through various agencies, it has launched programs and roadmaps to accelerate AI adoption. In particular, the Department of Science and Technology (DOST), through the Philippine Council for Industry, Energy, and Emerging Technology Research and Development (PCIEERD), published an AI roadmap to guide research and innovation, eventually evolving into the National AI Strategy Roadmap of the Philippines (NAIS-PH) in 2025.<sup>2</sup> The roadmap focuses on workforce upskilling and the Philippine government's thrust to educate learners to harness AI.

In 2024, the Department of Information and Communications Technology (DICT), together with the Civil Service Commission (CSC), organized a workshop to institutionalize a Joint Memorandum Circular (JMC) on the "Principles and Guidelines for an Ethical and Trustworthy Use of Artificial Intelligence (AI) in the Government." The JMC aims to ensure the responsible use, development, and deployment of AI across Philippine government agencies, grounded in the principles of human rights, transparency, data privacy, fairness, and accountability. Further, the JMC is expected to be aligned with international frameworks such as those established by the ASEAN and the OECD.

The Technical Education and Skills Development Authority (TESDA) has rolled out digital skills training, including courses on machine learning, data annotation, and cloud computing. TESDA's Learning Management System now hosts AI modules for the youth and professionals alike. Meanwhile, the Commission on Higher Education (CHED) supports curriculum innovation and AI faculty development through grants and partnerships with local universities.

In February 2025, the Department of Economy, Planning, and Development (DEPDev), formerly known as the National Economic and Development Authority (NEDA), released a landmark "Policy Note on Artificial Intelligence." The Policy Note calls for a "whole-of-nation" approach to fully leverage AI's impact on productivity, employment, public service

delivery, and economic growth. It recommends forming a National AI Ecosystem Council and aligning AI initiatives and programs with the Philippine Development Plan (PDP) 2023–2028.<sup>3</sup>

To this end, President Ferdinand Marcos Jr. has authorized the rollout of the National AI Upskilling Program in June 2025 following the recommendation of the Private Sector Advisory Council (PSAC).<sup>4</sup> The program will be led by CHED and TESDA, together with industry partners and stakeholders such as the Analytics and AI Association of the Philippines (AAP). The framework aims to equip Filipinos with the knowledge and skills to utilize AI, starting with the goal of educating 50,000 learners in two years.

It is evident in these policy initiatives that the Philippines' approach to AI adoption is evolving beyond a series of isolated projects into a broad-based ecosystem strategy—one that connects government, industry, academia, and startups in a dynamic loop of policy reform, capacity-building, and innovation. To this end, the Philippine Government must collaborate with—and build upon—the work of academia and the private sector.



### 2 Academia: driving research and talent development

Philippine universities are vital engines of AI talent and knowledge. The University of the Philippines (UP), Ateneo de Manila University (ADMU), and De La Salle University (DLSU), among others, have launched data science and AI degree programs and established research labs focused on real-world challenges, from agriculture and health to disaster response. In this regard, the academic sector is essential not only in producing technical talent and expertise but also in generating foundational research that is needed to sustain long-term AI advancement. These efforts can be further strengthened through international partnerships and by adopting interdisciplinary approaches which accelerate knowledge transfer and innovation.

### 3 The private sector: accelerating adoption and innovation

AI is also reshaping how businesses operate and compete in the global market. For the Philippines, the Information Technology and Business Process Management (IT-BPM) industry—employing over 1.3 million people—has been both challenged and catalyzed by AI. Instead of resisting change, leading business process outsourcing (BPO) companies are reinventing their service portfolios to offer AI support services such as chatbot development, sentiment analysis, data annotation, and analytics process outsourcing (APO).

IT-BPM companies are also retraining their workforce to adapt to new roles such as AI model validators, automation designers, model evaluators, data annotators, and customer experience analysts. Firms like Accenture, Concentrix, and TaskUs have partnered with educational institutions and global platforms like Coursera and Udemy to offer certification tracks in AI and data analytics.

Beyond the IT-BPM industry, AI is also gaining ground in healthcare (telemedicine triage, diagnostics), banking (fraud detection, credit scoring), logistics (predictive delivery systems), and retail (demand forecasting, personalized marketing). Philippine startups like Senti AI, Packworks, AI4Gov, and Rumarocket are building solutions for local needs—from language processing in Filipino dialects to agricultural forecasting. Innovation hubs like IdeaSpace and QBO Innovation Hub also provide critical support toward AI adoption. However, many SMEs lack the resources to adopt and leverage AI technologies, highlighting the need for scalable AI-as-a-service platforms and public support and investments for digital transformation. To this end, industry associations like the Analytics and AI Association of the Philippines (AAP) can play a central role. AAP can assist in building bridges between business, the government, and academia, and conduct capacity-building programs, policy dialogues, and research to advance the inclusive adoption and promote the responsible use of AI.

### Challenges in the Philippine AI ecosystem

Despite growing interest and notable foundational work, the Philippine AI ecosystem faces significant

challenges in infrastructure, human capital and education, funding and investments, as well as ethical and regulatory governance that must be addressed systematically.

#### 1 Infrastructure gaps

While urban centers enjoy internet connectivity and access to computing infrastructure, many rural areas lag behind. While government efforts such as the Free Wi-Fi for All Program and the National Broadband Plan are in motion, deployment is slow and often limited in scope. Data center infrastructure is also limited and the development of hyperscale and AI-ready data centers in the Philippines is constrained by high energy costs, unreliable power supply in certain regions, and limited renewable energy integration. These limitations hinder the ability to support large-scale AI training and cloud services, reducing competitiveness compared to neighboring countries with more robust energy and infrastructure ecosystems. Without equitable access to ICT infrastructure, AI adoption will remain uneven and confined to urban centers.

#### 2 Human capital and education challenges

There is a shortage of skilled AI practitioners, educators, and researchers in the Philippines. Universities are beginning to fill the gap, but demand far exceeds supply. Furthermore, much of the existing talent is absorbed by large tech firms abroad, contributing to local brain drain. Limited research grants and lab infrastructure for research development also remain a challenge. DOST-PCIEERD and other government agencies may provide an opportunity to address this challenge.

Public understanding of AI rights and responsibilities is also low, increasing risks of misuse or underuse of AI tools. As such, the Department of Education's (DepEd) E-Center for AI Research also needs further support to identify and implement plans for AI education at the basic level.

Resistance to AI adoption in traditional sectors like manufacturing, agriculture, and semiconductor industries can also present formidable barriers. The AAP, together with the Philippine Chamber of Commerce and Industry, remains in partnership to spread awareness on the importance of AI in non-IT-BPM and non-tech industries.

### 3 Funding and investment constraints

AI startups and university researchers struggle to secure long-term funding. Access to venture capital, GPU infrastructure, and commercial partnerships remains a constraint. Many startups lack long-term business models which could drive better identification of use cases. Non-government and civil society organizations also lack funding to scale training and advocacy programs for AI adoption. Without investment and incubation, many promising solutions remain at the pilot stage.

### 4 Ethical and regulatory barriers

AI has raised concerns about data privacy, algorithmic bias, and transparency. The Data Privacy Act of 2012 provides a foundation for data privacy protection in the Philippines, but regulations specific to AI are lacking.

Despite various government programs and private initiatives, coordination and alignment remain weak, leading to duplicated efforts and inefficiencies. Still, there is awareness of the issues as DEPDev's Policy Note calls for a National AI Ethics and Governance Framework. However, several draft bills currently filed in Congress, such as the "Proposed Artificial Intelligence Regulation Act," are still pending.

## A blueprint for collaboration

The Philippines is at a critical juncture in its AI journey. Despite challenges, numerous examples demonstrate the transformative power of AI when aligned with local priorities. The following real-world use cases illustrate the Philippines' progress in embedding AI in workforce development while addressing local needs in different industries and sectors:

1. **AI-driven skills mapping by DICT & TESDA:** Through AI-powered platforms, the DICT and TESDA are working to map existing and emerging digital skills gaps by region and industry, allowing for targeted training interventions and policy adjustments.
2. **AI for disaster resilience:** Local governments in partnership with research institutions are deploying AI tools for real-time flood forecasting, landslide prediction, and crisis

mapping, enabling frontline workers to make faster, data-driven decisions.

3. **Precision agriculture & farmer empowerment:** Startups like Cropital and other tech providers are training farmers to use AI apps for pest detection, yield forecasting, and weather prediction, integrating these tools into TESDA's agricultural training modules.
4. **BPO workforce reskilling with generative AI:** Companies like TaskUs are integrating generative AI tools into customer support workflows and upskilling employees to work alongside these tools in functions like content moderation and market research.
5. **Healthcare AI adoption in public hospitals:** The DOST's AI-assisted radiology projects train technicians and radiologists to work with diagnostic algorithms, improving healthcare access in underserved provinces.

In the coming years, greater integration of AI is expected in education, healthcare, public services, and commerce. These cases show the potential of AI to create new types of jobs, augment existing professions, and make public services more efficient. As AI and digital transformation deepens, the demand for AI-skilled professionals will rise, requiring continuous learning and agile curriculum development. With the right ecosystem in place, the Philippines could position itself as a regional leader in AI services, research, and workforce readiness, and emerge as a competitive player in the global AI economy. However, scaling these efforts requires coordinated public-private partnerships, continuous capacity-building, and adaptive policies. Strategic alignment among stakeholders—government, academia, industry, and startups—will be essential to ensure inclusive growth and widespread benefits. In this connection, the recommendations below identify areas where targeted collaboration can drive systemic improvements for the Philippine AI ecosystem.

1. **Operationalize DEPDev's recommendation for an AI Ecosystem Council.** A multi-sectoral council should serve as the central coordination body for national AI strategy, investments, and governance, bringing together various stakeholders from government, industry, academia, civil society,

and international partners. This council is vital to help build a robust, sustainable, and innovation-driven Philippine AI ecosystem.

2. **Strengthen regional infrastructure.** The government and other critical stakeholders should expand access to broadband, cloud credits, and computing facilities to support AI education and entrepreneurship beyond urban areas. Tapping renewable energy sources for AI infrastructure should also be encouraged and incentivized to help address energy supply gaps and costs.
3. **Invest in AI education and research.** CHED and TESDA can work with industry partners such as the AAP in implementing the approved National AI Upskilling Program and should continue supporting AI course development, faculty training, and industry-based internships.
4. **Support inclusive innovation through AAP and incubators.** The government and the private sector can scale up support for startups and SMEs through AAP, the Department of Trade and Industry's (DTI) QBO Innovation Hub, and university tech incubators, while also focusing on the social impact of the use cases. Industry associations like AAP have also been instrumental in mobilizing businesses, government, and academia to embrace data-driven transformation to bridge the skills and adoption gaps. AAP leads advocacy programs, capability-building workshops, and thought leadership forums that promote responsible and inclusive use of analytics and AI across sectors.

5. **Develop and enforce an AI ethics framework.** It is necessary to build on the Data Privacy Act with specific AI guidelines and promote responsible use across sectors.
6. **Launch national AI public awareness and literacy campaigns.** Such initiatives are aimed at improving the public's understanding of AI's benefits, risks, and applications, which is essential to building trust and encouraging its broader adoption.

AI offers the Philippines a powerful lever for economic resilience and social development. While infrastructure gaps, limited talent, and fragmented coordination remain as obstacles, these are not insurmountable. Through strong public-private-academic collaboration—spearheaded by institutions like CHED, DICT, DEPDev, and AAP—the country can build an ethical, inclusive, and globally competitive AI ecosystem.

With the right investments, governance, and vision, AI can become a force for upward mobility, regional equity, and innovation in the Philippines. AI is more than a technological trend—it is a developmental imperative. The convergence of government resolve, academic curiosity, industry transformation, and startup innovation sets the stage for meaningful progress. By embracing AI as a strategic priority, investing in talent, fostering innovation, and building an integrated ecosystem, the Philippines can create a resilient, inclusive, and future-ready workforce. This is not just about catching up with the rest of the world—it is about leading the way in shaping a smarter, more equitable future of work and economic prosperity for all Filipinos.

---

## References

- 1 "Human Development Report 2025 A matter of choice: People and possibilities in the age of AI," *UNDP*, May 6, 2025, <https://hdr.undp.org/content/human-development-report-2025>.
- 2 "DOST BUILDS ON "AI" NATIONAL STRATEGY," *DOST*, May 2, 2025, <https://www.dost.gov.ph/knowledge-resources/news/86-2025-news/3989-dost-builds-on-ai-national-strategy.html>.
- 3 "Policy Note on Artificial Intelligence," *DEPDev*, February 7, 2025, <https://depdev.gov.ph/policy-note-on-artificial-intelligence/>.
- 4 Dexter Barro II, "Private sector calls for roadmap on AI upskilling," *Manila Bulletin*, June 27, 2025, <https://mb.com.ph/2025/06/27/private-sector-calls-for-roadmap-on-ai-upskilling>.



*Photo taken during a breakout session at the 2024 KAS Solutions Conference.*

# Reimagining the Future of Experiences: Exploring the Realities of AI Integration in the Philippine Tourism and Hospitality Industry

**Eylla Laire M. Gutierrez and Ray Justin A. Villanueva**

The tourism and hospitality (T&H) sectors have long served as vital pillars of the Philippine economy. The Philippine Statistics Authority (PSA) reported that the sector contributed 12.7% to the country's gross domestic product (GDP) in 2019, with an average GDP contribution of 7.4% from 2000 to 2018. Although the COVID-19 pandemic disrupted this positive trajectory, the sector has demonstrated resilience, with tourism contributing 8.6% to the country's GDP in 2023.<sup>1</sup> The sector also remains a key driver of employment, accounting for 34% of the national workforce, or approximately 16.4 million Filipinos in 2024.<sup>2</sup>

The sector's economic impact also extends well beyond major urban centers, making it a cornerstone of regional and local development. The widespread presence of small and medium enterprises (SMEs) engaged in tourism-related activities spanning from accommodation, transportation, food and beverage, and curated visitor experiences, among others, reflects a growing diversification of travel preferences and a broader geographic spread of tourism benefits across the country. As a result, more regions are able to engage and benefit from tourism-driven economic development.

These figures highlight the sector's crucial role as an economic driver and a source of income and livelihood for millions of Filipinos. Nevertheless, the country's reliance on tourism revenue underscores the urgent need

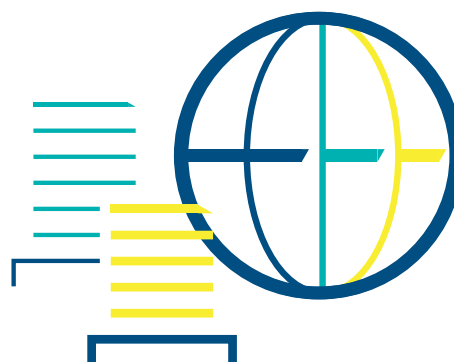
to ensure its adaptability and competitiveness, particularly in the context of rapid digitalization and the growing integration of artificial intelligence (AI) in the sector.

Globally, AI applications such as generative AI,<sup>3</sup> robotics,<sup>4</sup> chatbots and virtual assistants,<sup>5</sup> as well as predictive analytics and dynamic pricing algorithms are becoming standard tools that are revolutionizing many firms in the T&H sectors. These technologies have been instrumental in streamlining operations through process automation and big data analysis, allowing for more tailored marketing and improved customer experiences.<sup>6</sup> From the demand side, these tools assist travelers in accessing more relevant information, thereby improving decision-making and contributing to a richer and more satisfying experience.<sup>7</sup> While digitalization has been integral across all stages of a traveler's journey,<sup>8</sup> it is AI which enabled a more seamless integration of physical and virtual touchpoints.<sup>9</sup> It has created systems that bridge online platforms with on-ground realities, transforming how travelers engage with destinations and increasing the overall value of tourism experience. While this rapid pace of AI adoption presents benefits and opportunities, particularly for economies such as the Philippines that have a sizable T&H industry, challenges such as the digital divide and a general lack of awareness and understanding of AI within the business sector hinder its integration, including in tourism enterprises.<sup>10</sup>

Within the Philippine context, digital and AI integration within businesses remains in the nascent stage. According to the 2023 Asia Pacific AI Readiness Index, the Philippines ranked 12<sup>th</sup> out of 12 countries in terms of overall AI readiness as well as government AI readiness, and 10<sup>th</sup> on business AI readiness, signifying a substantial gap in the country's preparedness to adopt and harness AI technology.<sup>11</sup> This limited readiness is also evident in the Philippine T&H sector as AI adoption remains uneven and mainly concentrated on larger enterprises in the metropolitan area. In contrast, smaller-scale enterprises, particularly in rural areas, are lagging due to limited digital infrastructure, challenges in human capital development and awareness, and insufficient access to funding and capital needed for AI investments and broader digital transformation initiatives. Nevertheless, several tourism stakeholders in the country,

including the academe, industry practitioners, and government agencies have positive perceptions toward the 4<sup>th</sup> Industrial Revolution (4IR) and the role of technology in the industry.<sup>12</sup> However, issues such as the potential displacement of workers, algorithmic bias, and concerns on data privacy stagers the broader adoption of digitalization in the country, including in the T&H sector.

Given the vital contribution of tourism and hospitality to the Philippine economy and digitalization's transformative potential, it is essential to examine the dynamics of digital and AI adoption within the sector and provide practical insights that can guide policymakers, industry practitioners, and the academe in fostering a future-ready T&H sector in the Philippines.



## A global overview

The implementation of AI and digitalization in the T&H sector is transforming both the industry and customer experience. Its benefits have been well documented across literature, suggesting its profound impact on travelers' satisfaction, economic outcomes, and in supporting innovative business models that can transform the industry as a whole.<sup>13</sup> This movement, driven by the goal of further enhancing guest satisfaction, operational efficiency, and cost-saving measures, can be traced back to the 2010s where various AI-powered technological developments started penetrating the T&H sector.

Among these advancements are intelligent service robots such as "Mario"—a robot receptionist in the Ghent Marriott Hotel in Belgium; "Connie"—the robotic concierge operating in the Hilton McLean Hotel in the United States; "Spencer"—an airport escort operating in Amsterdam Schiphol Airport in

the Netherlands; and “Botlr”—the robot butler used in Aloft Hotels.<sup>14</sup>

AI digital voice assistants (VAs) and chatbots have also been widely integrated in the industry, mostly at large-scale infrastructure and businesses, such as airports and hotel chains, to enhance customer experience and satisfaction.

AI-powered VAs, which utilize voice recognition, machine learning, and natural language processing (NLP), act as digital assistants that can execute tasks or provide information through spoken commands or speech-enabled interactions. Research has found that digital VA technology can help hotels deliver better service while streamlining operations and costs.<sup>15</sup> For example, Marriott International is deploying AI-powered in-room assistants that respond to voice commands to adjust various room features such as lighting, entertainment systems, and temperature, creating a more personalized stay for their guests.<sup>16</sup>

In parallel, chatbots on reservation platforms such as Booking.com and Skyscanner, among others, have already been enhanced by the presence of AI assistants that allow travelers to book flights, accommodations, and transportation arrangements online.<sup>17</sup> Through AI, the booking process is streamlined according to traveler preference—from budget, duration, and level of comfort, among other factors. These chatbots are also trained on large datasets using AI and machine learning to generate human-like text or voice replies to respond to inquiries. AI chatbots on hotel websites and social media platforms have also facilitated improved customer service as these tools enabled instant responses to guest queries and concerns. The case of Edwardian Hotels’ “Edward” shows how AI chatbots can serve as a concierge, providing 24/7 assistance to guests in multiple languages. The hotel credits “Edward” for a 10% to 50% increase in room service sales between 2016 and 2019.<sup>18</sup>

The application of AI-powered solutions to support back-end functions is also prevalent across the industry. AI has been utilized to support inventory management in hotels, using data to forecast and manage stock levels. Others have also used AI in employment management to manage workflow and housekeeping schedules alongside hotel traffic data to ensure maximum efficiency.<sup>19</sup>

Processing and analyzing feedback from customers can also be improved by AI using tools or apps such as TrustYou that can sift through online reviews, conduct sentiment analysis, manage data, and automate responses to feedback and reviews. AI can also be used to deliver personalized experiences to guests even before check-in through targeted marketing. An example of this is AccorHotels, which leverages AI to analyze guest preferences and booking histories to provide recommendations that enhance guest loyalty.

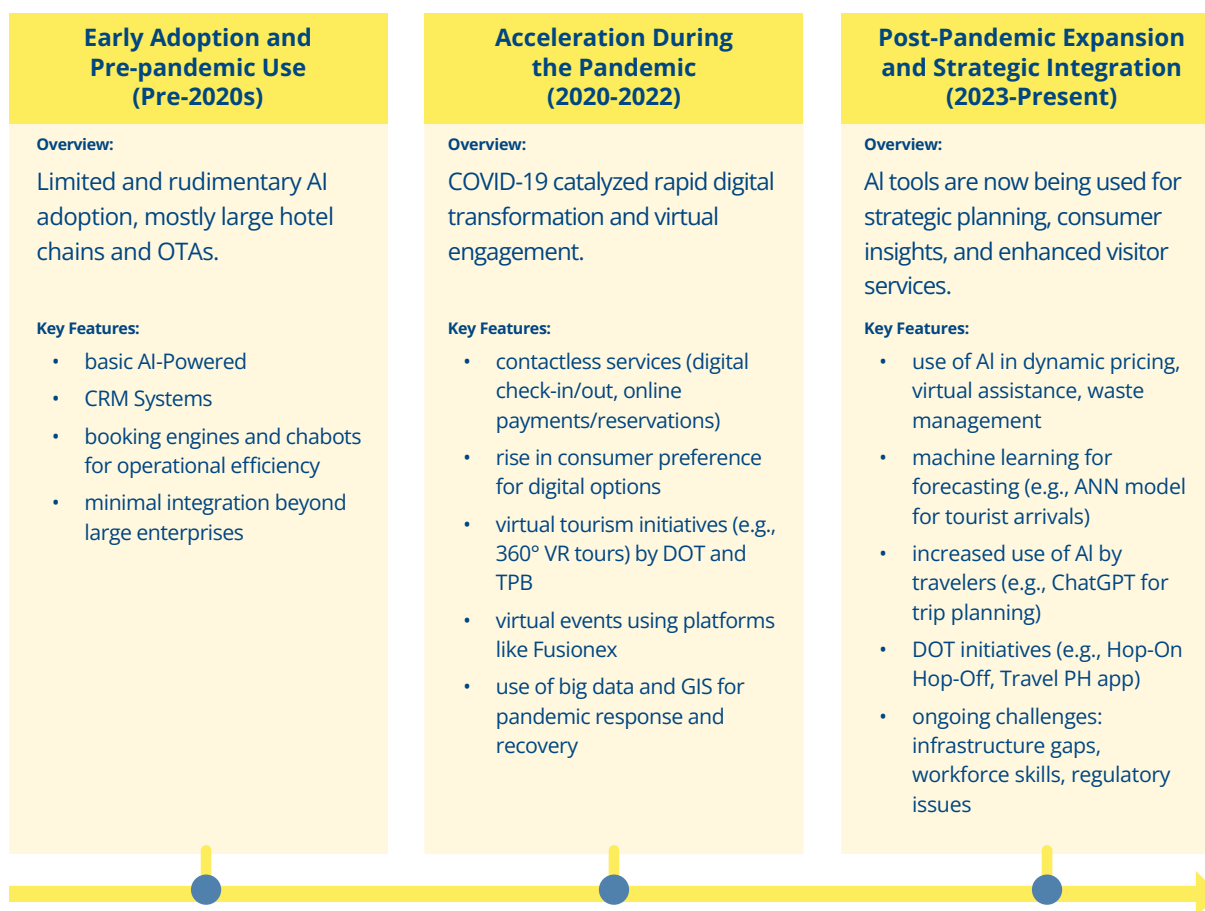
Within the restaurant industry, service automation has been emerging through the “robotics, artificial intelligence, and service automation” (RAISA) model. The model works in various ways, from kitchen automation, menu recommendation, customer service, marketing, cleaning, inventory management, and quality control.<sup>20</sup> This automation system is also now observable within the hotel industry.

Beyond these efforts, AI integration has also been observed to facilitate sustainability initiatives (e.g. energy management), enhance security measures (e.g. surveillance systems), develop new tourism products (e.g. virtual reality tours), improve customer experience (e.g. predicting room service), and promote inclusivity (e.g. AI-solutions assisting older or people with disabilities (PWDs) when travelling).<sup>21</sup> These various applications showcase how AI is reshaping the global T&H landscape by creating more personalized and sustainable experiences, enhancing accessibility, and driving efficiency across diverse guest touch points.

### **Zooming in: AI integration in Philippine tourism and hospitality**

Over the past decade, AI adoption has steadily gained traction in the Philippine T&H sector, following a distinct trajectory shaped by the need for technological advances, pandemic-driven necessity, and continuous efforts in innovation. This progression, as outlined in Figure 1, traces AI adoption within the Philippine T&H sector, from early attempts at integration to pandemic-driven acceleration, and more recently, to strategic integration. However, despite key milestones and notable best practices by several organizations, technological and workforce challenges, as well as ethical concerns, hinder the broader integration of AI in the industry.

**Figure 1. Timeline of AI Adoption in the Philippine T&H Sectors**



As shown in the first column of the diagram, AI adoption in the Philippine T&H sectors was relatively limited prior to the 2020s. Insights from various stakeholder consultations and sectoral dialogues on AI integration support that early applications had been largely rudimentary and focused on automating the operations of large hotel chains and online travel agencies (OTAs). Typical applications include the use of booking engines and AI-powered chatbots integrated with basic customer relationship management (CRM) systems primarily deployed to improve operational efficiency and customer service.

The COVID-19 pandemic marked a significant turning point for AI adoption. The implementation of travel restrictions, social distancing protocols, and hygiene measures prompted an accelerated shift to digitize services to ensure safety and business continuity.<sup>22</sup> The industry saw large hotel chains and other tourism establishments, especially in urban areas, implement contactless services such as digital check-in and check-out systems, cashless payments, and online reservations to

reduce physical interaction and ensure the safety of both guests and staff. Meanwhile, with strict restrictions and border closures in effect, businesses shifted to virtual engagements, using online platforms such as Zoom, Webex Events, Virtual Expo Philippines, and Fusionex to launch exhibitions, host webinars, and run campaigns. This shift improved accessibility and sustained customer engagement during a time of limited mobility.<sup>23</sup>

While not necessarily AI-driven, these initiatives have helped in laying the groundwork for increased openness and readiness to digital and intelligent systems among Filipinos. In fact, these changes were met with favorable reception. As noted in a joint survey released in 2020 by the Department of Tourism (DOT) and the Asian Institute of Management–Dr. Andrew L. Tan Center for Tourism (AIM-ALT CFT), more Filipinos have shown a growing preference for digital payments, online reservations, and self-service options even after the easing of leisure travel restrictions.<sup>24</sup> These broader shifts also helped build the T&H industry’s digital maturity and readiness for more advanced

technologies such as AI. However, reflected within documented cases and practices is a persistent gap in the industry's actual AI adoption. Most of these innovations remain focused on automation and improved service delivery rather than renewed, transformative AI-integrated system designs.

Nevertheless, AI integration in the tourism enterprises and even the public sector has continued to gain traction after the COVID-19 pandemic. These enterprises and agencies have leveraged the use of AI tools not only to improve customer service but also to harness data-driven insights on emerging trends and consumer behavior, thereby supporting business decisions and sustainability, as well as guiding development efforts. Major hotel chains across the country have also begun increasingly deploying AI-integrated smart hotel systems such as dynamic pricing algorithms, targeted marketing, virtual assistance technology, intelligent surveillance, behavior monitoring, and guest safety systems,<sup>25</sup> as well as AI-driven waste management systems.<sup>26</sup>

In terms of strategic integration, Baguio City has taken initial steps toward becoming a smart city by leveraging data science and AI in addressing urban planning challenges and promoting sustainable development, including by managing tourism demands and analyzing key points of interest within the city through its Project MINERVA. There is also evidence from research that neural networks, as predictive analysis tools, can be effective in forecasting tourist arrivals in the Philippines.<sup>27</sup> Meanwhile, generative AI chatbots, such as ChatGPT, have also been increasingly used by Filipino travelers for planning personalized itineraries.

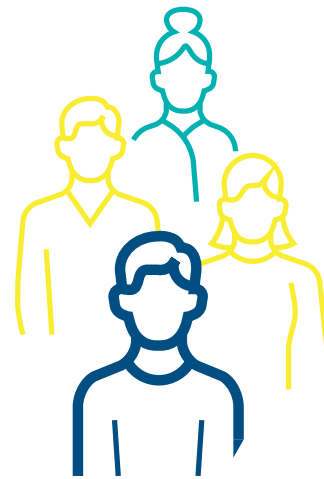
Despite these significant developments, the broader adoption of AI in the Philippine T&H sector remains constrained by critical barriers, including technological limitations, workforce capability gaps, and ethical and regulatory concerns.

## Barriers and key challenges

### Technological challenges

One of the primary barriers to AI adoption in the Philippines is the lack of consistent and reliable ICT infrastructure, particularly in rural and tourism-dependent areas.<sup>28</sup> This persistent digital divide

limits the ability of enterprises and local institutions to adopt and scale AI-enabled platforms effectively. This disparity is worsened by limited access to funding and investment opportunities, particularly in the Philippine AI landscape, where high implementation costs<sup>29</sup> disproportionately favor large-scale enterprises with the resources to invest in these highly advanced tools and systems. Issues with data accessibility and limited information on the use of high-quality training datasets, which can lead to incorrect data application and model development, are also cited as challenges in the Philippine AI landscape.<sup>30</sup> As a result, many SMEs, both within and outside the T&H sectors, are left without the means to leverage the potential of AI in improving efficiency, marketing, and customer engagement.



### Workforce challenges

Effective AI adoption requires a digitally literate and future-ready workforce. However, in the Philippines, the general lack of awareness and skills mismatches, particularly in rural areas, creates a bottleneck in adoption and widens the gap between tech-enabled enterprises and traditional service providers. A recent study of the Philippine Institute of Development Studies (PIDS) reveals that only one in five firms demonstrates familiarity with AI technologies, highlighting a significant knowledge gap in the country.<sup>31</sup>

Compounding this challenge is the prevailing fear of job displacement. According to the recent report of the International Monetary Fund (IMF),<sup>32</sup> 36% of the jobs in the Philippines are significantly impacted by AI, with 14% at direct risk of replacement. These include roles in the business process outsourcing

(BPO) sector, clerical work, and sales—many of which are linked, either directly or indirectly, to the T&H sectors—in positions such as frontline workers and support staff within hotels and related enterprises. While AI presents opportunities to enhance productivity and create new roles, its successful integration necessitates comprehensive upskilling and reskilling programs, requiring significant investments in both financial and technical resources.

### **Ethical and regulatory challenges**

Despite ongoing efforts to establish legal frameworks and roadmaps supporting AI usage and development in the country, ethical and regulatory challenges continue to present formidable barriers. More specifically, the possible misuse of personal information by government and private entities to illegally or unfairly track and discriminate individuals have been pointed out,<sup>33</sup> coupled by concerns on the possibility of such information to be used for surveillance purposes.<sup>34</sup> Within the T&H industry, algorithmic biases and manipulation of consumer demand were also cited to be one of the challenges posed by AI, especially where predictive analytics easily influences consumer choices and interests.<sup>35</sup> Cultural resistance and public apprehension towards AI adoption, often rooted from the fear of the unknown, complemented by the perceived threats on job displacement, and concerns on trust towards automated systems can also pose additional barriers. These concerns are particularly relevant, especially in tourism and hospitality where personal and human-driven service is highly valued. Although information on these perceptions remains limited due to insufficient literature, these emerging sentiments have been noted in broader discussions with various stakeholders and key industry players regarding AI integration. Moreover, these concerns are further compounded by the absence of comprehensive data protection guidelines and ethical frameworks on the proper use of AI in the workplace and the T&H industry.

A recent survey of Deloitte Philippines on generative AI (gen AI) adoption, which included over 2,800 C-suite and director-level respondents from six industries across 16 countries, including the Philippines, reveals that 36% of business leaders expressed a lack of confidence on results generated by AI, 35% raised intellectual property concerns,

and 33% reported uncertainty about their ability to comply with regulations.<sup>36</sup> Another survey of 35 Philippine-based CEOs released by PwC in 2024 found that while AI is seen as a driver of efficiency, revenue, and profitability, 69% believe that it will increase cybersecurity risks and 57% are concerned that it could even be a tool for misinformation, exacerbating this existing problem in the Philippines.<sup>37</sup> Although these findings are not exclusive to the T&H industry, it is indicative of the broader sentiment, especially industry reluctance on adopting AI technologies.

Without clear regulatory frameworks and robust ethical guidelines and safeguards in place, these apprehensions are more likely to persist limiting the potential for a more inclusive and socially responsible AI adoption in the country, particularly in the T&H sectors where service quality and delivery significantly relies on authentic, personalized, and trust-based interactions that AI may struggle to replicate without risking possible depersonalization or ethical breaches.

### **Toward a more inclusive and ethical AI integration**

The adoption of AI in the Philippine T&H industry often points to a gray area where trade-offs, effects, and contradictions may be observed. As a result, parameters that ensure inclusivity and ethical considerations are paramount to its implementation. Stakeholder engagement and collaboration across sectors prove to be crucial in building trust, addressing potential risks, and ensuring inclusivity in the process of AI integration in the T&H sector.<sup>38</sup> Collaborative design processes and mechanisms, in the form of AI tourism innovation hubs and public-private partnerships (PPPs) among others, may be put in place to enable a participatory approach to AI integration processes. Engaging industry stakeholders, especially those considered in vulnerable positions, is key to this.<sup>39</sup> Policymakers, private industry stakeholders, and academia need to work together closely in making this happen. Through such collaborative work, needs-based interventions may be promoted. In this connection, the authors propose a three-pronged recommendation that hinges on stakeholder engagement and cross-sector collaboration on educational and training programs, infrastructure investments, and ethical framework development.

### 1 Educational and training programs

As AI solutions become increasingly embedded in the Philippine T&H sector and broader society, awareness campaigns now need to be coupled with educational and training programs that specifically cater to AI readiness defined as “a contextualized way of helping people to understand artificial intelligence.”<sup>40</sup> While AI awareness campaigns that highlight the benefits of AI and address misconceptions are being done in general, efforts to specifically ensure the AI readiness of workers, business owners, managers, and stakeholders involved in the T&H sector also need to be rolled out. Beyond knowing what AI means, education also needs to expand toward helping people understand how AI can meet their needs and strengthen their respective competitive advantages.<sup>41</sup> This is particularly important in addressing concerns about AI’s capacity to displace workers. With such programs in place, stakeholders from the T&H sector can work more proactively, rather than hesitantly, as the move towards AI integration becomes inevitable.

### 2 Infrastructure investment

Harnessing the projected benefits of AI requires infrastructure capacity that is encapsulated within the idea of smart tourism models. Smart tourism has become a primary investment point where ICT infrastructures are integrated to create improved tourism experiences. Investments for infrastructure development, therefore, go beyond providing reliable internet access to creating an integrated system in which technology serves travelers from planning to the actual visit itself. This investment approach is particularly important in the Philippine T&H context, where several of its stakeholders, including SMEs, are situated outside of Metro Manila and in rural areas where digital infrastructure remains limited. As such, need-based infrastructure development, where stakeholders’ insights, interests, and requirements are accounted for, is essential.

### 3 Ethical framework development

AI integration into the Philippine T&H sector needs to be done responsibly and fairly, given that AI technologies are becoming more embedded in the industry, thus heightening the risks of ethical breaches and privacy violations. To address these

issues, cooperation between the government and private industry members is necessary to establish ethical frameworks that provide guidelines on the use of AI in the industry and ensure that these are aligned with human rights and societal values. In doing so, fairness, transparency, accountability, and human oversight remain at the core of AI systems.<sup>42</sup> Guidelines should also include provisions for ensuring fairness-aware algorithms, regular audits of AI systems, data privacy regulations, and accountability measures. A T&H sector-specific set of ethical guidelines may borrow from existing frameworks such as the OECD AI Principles and the Montreal Declaration for a Responsible Development of Artificial Intelligence, among others.



Ray Justin Villanueva (leftmost) and Eylla Laire Gutierrez (rightmost) during the 2024 KAS Solutions Conference.

### Moving forward

AI, particularly in the Philippine T&H sector, has evolved beyond being a trend to become a paradigm shift, setting new and better standards for exceptional service and guest experiences in this digital age. However, much of the innovation remains concentrated on automation and incremental improvements in service delivery, particularly in urban or high-traffic areas. Significant gaps still exist in terms of infrastructure, workforce readiness, and ethical safeguards, limiting the full potential of AI to drive inclusive, transformative change across the broader tourism and hospitality ecosystem. The challenge remains in the balancing act required to preserve human interaction while facilitating technological innovation.

As some have pointed out, technologies such as AI that carry out tasks previously done by humans are not intended to replace entry-level jobs. Instead, AI can be seen as a complement to human touch, amplifying service quality, not replacing it. There is

no replacement for quality customer service provided by humans.<sup>43</sup> When implemented responsibly, AI can enhance—not diminish—the uniquely human experiences that define Philippine tourism and hospitality.

Moving forward, unlocking the true value of AI for the Philippine T&H sector demands a holistic

approach that addresses structural challenges while keeping human connection at the heart of innovation. This emerging challenge and opportunity underscore the urgent need for the country to bridge digital divides, empower stakeholders, and foster ethical, inclusive, and sustainable adoption of AI so that no one is left behind in the future of travel.

## References

- 1 "Share of Tourism to GDP Is 12.7 Percent in 2019" Philippine Statistics Authority, June 19, 2020, <https://psa.gov.ph/statistics/tourism/node/162606>.
- 2 Joyce Ann L. Rocamora, "DOT: Equitable tourism development factor in growth of PH regions," *Philippine News Agency*, April 25, 2025, <https://www.pna.gov.ph/articles/1248747>.
- 3 Galina Ilieva, Tania Yankova, and Stanislava Klisarova-Belcheva, "Effects of Generative AI in Tourism Industry," *Information* 15, no. 11 (2024): 671, <https://doi.org/10.3390/info15110671>.
- 4 Iis Tussyadiah, "A review of research into automation in tourism: Launching the Annals of Tourism Research Curated Collection on Artificial Intelligence and Robotics in Tourism," *Annals of Tourism Research* 81, (2020): 102883, <https://doi.org/10.1016/j.annals.2020.102883>.  
Sangwon Park, "Multifaceted trust in tourism service robots," *Annals of Tourism Research* 81, (2020): 102888, <https://doi.org/10.1016/j.annals.2020.102888>.
- 5 Zohreh Dobarjeh, Nigel Hemmington, Maryam Dobarjeh, and Nikola Kasabov, "Artificial intelligence: a systematic review of methods and applications in hospitality and tourism," *International Journal of Contemporary Hospitality Management* 34, no. 3 (2022): 1154-1176, <https://doi.org/10.1108/ijchm-06-2021-0767>.
- 6 Seden Doğan and İlayda Zeynep Niyet, 2024. "Chapter 1: Artificial Intelligence (AI) in Tourism," in *Future Tourism Trends Volume 2: Technology Advancement, Trends and Innovations for the Future in Tourism*, edited by Canan Tanrisever, Hüseyin Pamukçu, and Anukrati Sharma (Emerald Publishing Limited, 2024), 3–21, <https://doi.org/10.1108/978-1-83753-970-320241001>.
- 7 Lukas Grundner and Barbara Neuhofer, "The bright and dark sides of artificial intelligence: A futures perspective on tourist destination experiences," *Journal of Destination Marketing & Management* 19, (2021): 100511, <https://doi.org/10.1016/j.jdmm.2020.100511>.
- 8 Ibid.
- 9 Mouna Knani, Said Echchakoui, and Riadh Ladhari, "Artificial intelligence in tourism and hospitality: Bibliometric analysis and research agenda," *International Journal of Hospitality Management* 107, (2022): 103317, <https://doi.org/10.1016/j.ijhm.2022.103317>.
- 10 Francis Mark A. Quimba, Neil Irwin S. Moreno, and Alliah Mae C. Salazar, "Readiness for AI adoption of Philippine Business and Industry: The Government's Role in Fostering Innovation- and AI-Driven Industrial Development," *Philippine Institute for Development Studies Discussion Paper Series*, December 19, 2024, <https://doi.org/10.62986/dp2024.35>.
- 11 "New Salesforce Report Reveals How Asia Pacific AI Readiness Will Impact Deployment of Generative AI," *Salesforce*, September 28, 2023, <https://www.salesforce.com/news/stories/new-salesforce-report-reveals-how-asia-pacific-ai-readiness-will-impact-deployment-of-generative-ai/>.
- 12 Pia Rhoda Pinpin-Lucero, "Perception of Stakeholders on the Adoption of 4th Industrial Revolution Technologies in the Hospitality Industry in the Philippines," in *Tourism in the Philippines*, edited by Richard S. Aquino and Brooke A. Porter (Springer, 2022), 151–66, [https://doi.org/10.1007/978-981-19-4497-0\\_8](https://doi.org/10.1007/978-981-19-4497-0_8).
- 13 Edward C.S. Ku, and Chun-Der Chen, "Artificial intelligence innovation of tourism businesses: From satisfied tourists to continued service usage intention," *International Journal of Information Management* 76, (2024): 102757, <https://doi.org/10.1016/j.ijinfomgt.2024.102757>.
- 14 Natalie Holmes, "Mario, the robot receptionist, speaks 19 languages and dances to Michael Jackson," *World Economic Forum*, June 28, 2016, <https://www.weforum.org/stories/2016/06/mario-the-robot-receptionist-speaks-19-languages-and-dances-to-michael-jackson>.  
Dimitrios Buhalis and Luliia Moldavska, "Voice assistants in hospitality: using artificial intelligence for customer service," *Journal of Hospitality and Tourism Technology* 13, no. 3 (2022): 386-403, <https://doi.org/10.1108/JHTT-03-2021-0104>.
- 15 Neil Sahota, "AI in Hospitality: Elevating the Hotel Guest Experience through Innovation," *Forbes*, March 6, 2024, <https://www.forbes.com/sites/neilsahota/2024/03/06/ai-in-hospitality-elevating-the-hotel-guest-experience-through-innovation/>.
- 17 Or Eldar Izchak, "6 Examples of How AI is Used in the Travel Industry," *mize*, April 22, 2025, <https://mize.tech/blog/6-examples-of-how-ai-is-used-in-the-travel-industry/>.
- 18 Dorothy Creamer, "Chatbot Increases Hotels' Room Service Sales 10-50% and Gets Wooed by Guests," *Hospitality Technology*, December 13, 2019, <https://hospitalitytech.com/chatbot-increases-hotels-room-service-sales-10-50-and-gets-wooded-guests>.

- 19 Sahota, "AI in Hospitality."
- 20 Joanna Santiago, Maria Teresa Borges-Tiago, and Flávio Tiago, "Embracing RAISA in restaurants: Exploring customer attitudes toward robot adoption," *Technological Forecasting and Social Change* 199, (2024): 123047, <https://doi.org/10.1016/j.techfore.2023.123047>.
- 21 Kangping Wang, Haiyan Kong, Naipeng Bu, Honggen Xiao, Xuejie Qiu, & Jinghan Lim, "AI in health tourism: developing a measurement scale," *Asia Pacific Journal of Tourism Research* 27, no. 9 (2022): 954-966, <https://doi.org/10.1080/10941665.2022.2142620>.
- 22 "Philippine Tourism Sector Adopts Virtual Platform for Business Continuity," *OpenGov Asia*, August 23, 2021, <https://opengovasia.com/2021/08/23/philippine-tourism-sector-adopts-virtual-platform-for-business-continuity/>.
- 23 Ibid.
- 24 "Domestic travel to drive recovery of tourism industry," *Philippine News Agency*, June 29, 2020, <https://www.pna.gov.ph/articles/1107350>.
- 25 Tom Chapman, "How Okada Manila is taking full advantage of the AI boom," *Technology Magazine*, October 6, 2023, <https://technologymagazine.com/company-reports/how-okada-manila-is-taking-full-advantage-of-the-ai-boom>.
- 26 Claudia Light, "How a Philippines hotel cut food waste by 67% with AI," *Winnow*, August 9, 2024, <https://blog.winnowsolutions.com/swissotel-clark-philippines-cuts-food-waste-by-67-percent>.
- 27 Noelyn M. De Jesus and Benjie R. Samonte, "AI in Tourism: Leveraging Machine Learning in Predicting Tourist Arrivals in Philippines Using Artificial Neural Network," *International Journal of Advanced Computer Science and Applications* 14, no. 3 (2023): 816-823, <https://doi.org/10.14569/ijacsa.2023.0140393>.
- 28 Quimba, Moreno, and Salazar, "Readiness for AI adoption of Philippine business and industry."
- 29 Nelidiza R. Arceta, Marvin C. Montero, and John Emmanuel P. Quilitis, "Adaptation and Effectiveness of Artificial Technology Used by Selected Hotels in Quezon City," *International Journal of Tourism and Hospitality* 4, no. 2 (2024): 27-46, <https://doi.org/10.51483/IJTH.4.2.2024.27-46>.
- 30 Ibid.
- 31 Quimba, Moreno, and Salazar, "Readiness for AI adoption of Philippine business and industry."
- 32 "Philippines: 2024 Article IV Consultation-Press Release; and Staff Report," *IMF Staff Country Reports* 2024, no. 351 (2024): 23, <https://doi.org/10.5089/9798400296246.002>.
- 33 Brenda Leong, "Facial recognition and the future of privacy: I always feel like... somebody's watching me," *Bulletin of the Atomic Scientists* 75, no. 3 (2019): 109-115, <https://doi.org/10.1080/00963402.2019.1604886>.
- 34 Shoshana Zuboff, *The age of surveillance capitalism in Social theory re-wired* (Routledge, 2023), 203-213.
- 35 Stefan Gössling, "Extending the theoretical grounding of mobilities research: Transport psychology perspectives," *Mobilities* 18, no. 2 (2023): 167-183, <https://doi.org/10.1080/17450101.2022.2092886>.
- 36 "New Deloitte Survey on Gen AI Adoption," *Deloitte*, February 28, 2024, <https://www.deloitte.com/southeast-asia/en/Industries/tmt/perspectives/now-decides-next.html>.
- 37 "Asia Pacific CEO Survey 2024 (Philippines)," *PwC*, June 26, 2024, <https://www.pwc.com/gx/en/asia-pacific/ceo-survey/27th-ceo-survey-asiapacific-philippines.pdf>.
- 38 Olivia J. Erdélyi and Judy Goldsmith, "Regulating Artificial Intelligence: Proposal for a Global Solution," *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, December 27, 2018, 95-101, <https://doi.org/10.1145/3278721.3278731>.
- 39 Dimitrios Sargiotis, "Fostering Ethical and Inclusive AI: A Human-Centric Paradigm for Social Impact," SSRN, July 29, 2024, <https://ssrn.com/abstract=4879372>.
- 40 Rose Luckin, Jürgen Rudolph, Martin Grünert, and Shannon Tan, "Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin," *Journal of Applied Learning and Teaching* 7, no. 1 (2024): 346-363, <https://doi.org/10.37074/jalt.2024.7.1.27>.
- 41 Ibid.
- 42 "Ethics guidelines for trustworthy AI," *European Commission*, April 8, 2019, <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- 43 "Robot butlers 'a gimmick', says hotel workers' union," *BBC*, August 18, 2014, <https://www.bbc.com/news/technology-28834773>.



# Harnessing AI for Good Governance

*Marie Antoinette P. de Jesus during the 2024 KAS Solutions Conference.*

## Artificial Intelligence in Campaigns and Elections: A Force for Empowerment or a Threat to Democracy?

**Marie Antoinette P. de Jesus**

---

The growth of artificial intelligence (AI) has resulted in rapid changes to the way political campaigns are conducted and the way citizens take part in elections. In recent years, AI tools have been deployed in election campaigns worldwide, raising hopes for more efficient and personalized voter outreach, while also raising concerns over new forms of manipulation. The very different uses of AI raise a key question: Is AI helping support democratic electoral participation, or is it putting it at serious risk?

In 2024, about half of the world's population, roughly 3.7 billion people across 72 nations, were eligible to vote.<sup>1</sup> Throughout this year of major global elections, AI has been subtly, but significantly changing how democracy works around the world. Given this context, it is imperative to look at both sides of AI's role in elections. On the one hand, AI can help improve political campaigns and raise voter engagement through personalized messages and automatic fact-checking. On the other hand, it also threatens fair elections by spreading false information, creating fake videos and audio, and manipulating voters through microtargeted abuses.<sup>2</sup> Lessons learned from real-world examples show the need for rules and guidelines that help maximize the benefits of AI while keeping democratic systems safe and trustworthy.

### Transforming campaigns and voter engagement with AI

Advocates believe AI can considerably increase legitimate political participation by making the playing field more equal and helping campaigns get in touch with voters more efficiently and effectively than ever before.

In the U.S., political groups have started using AI to speed up their fundraising processes. In 2024, the nonprofit Tech for Campaigns tested AI-generated fundraising emails for state-level candidates and found that messages written with help from OpenAI's ChatGPT brought in 4.4 times more money per hour of work.<sup>3</sup>

Even with a human team still reviewing and editing the content, these experiments prove how AI can save time and resources, letting campaign staff focus more on connecting with voters and planning strategy. Moreover, even if AI just matches what humans can do in writing, it still helps shift effort and resources toward more personal voter outreach and campaign work.<sup>4</sup>

The efficiency gains from AI can level the playing field for smaller, underfunded campaigns. Although they cannot afford the most advanced or expensive tools, access to basic AI technologies still gives them something to work with, a way to somehow compete against better-resourced opponents. This is particularly relevant in countries like the Philippines, where candidates typically need substantial funding to run winning campaigns.



Another way AI is being used in elections is through personalized and highly targeted messaging. While campaigns have always tried to customize their messages for different voter groups, AI now allows them to do this instantly and with much more precision. Today's AI tools can analyze voter data, like demographics, voting history, and social media activity, among others, to profile the electorate and spotlight what different groups care about. Candidates can then craft messages or policy pledges that are targeted or more responsive to the needs and priorities of their constituents. The goal is not for candidates to fabricate concern over an issue, but rather to convince voters by highlighting how political aspirants and their platforms can meaningfully address priorities and challenges. When tailored communication is carried out in a way that is authentic, transparent, and ethical, voters will feel genuinely heard, deepening their engagement in the political process.

For the Philippines, data experts note that campaigns could use AI to assess public

perceptions and even run their own virtual surveys to determine where they can secure votes from the population.<sup>5</sup> Filipino AI scientist Allan Tan observes that AI lets candidates conduct independent surveys and apply algorithmic analyses on voter data rather than relying solely on traditional polling.<sup>6</sup>

AI helps campaigns reach wider audiences and also makes their messages more accessible. In multilingual countries, AI translation tools help campaigns engage minority groups in their own languages. Research in 2024 showed that advanced AI systems can now translate text and speech instantly, cutting costs and giving politicians access to voters they previously could not reach.<sup>7</sup>

AI can also instantly generate versions of pre-recorded speech in multiple local languages, even maintaining the speaker's voice and intonation. This was a real case during the 2020 elections in Delhi, India, where one political party released an AI-generated video of its candidate speaking in a dialect that the candidate did not actually speak in order to connect with voters. The video is essentially a deepfake. However, the material was clearly labeled as a campaign message and used for outreach to a marginalized language community.<sup>8</sup>

These examples show the capability of AI to create opportunities to make politics more accessible through locally customized messaging, instant translation, audio features for people with visual impairments, and captions for those with hearing difficulties. These tools help remove language and accessibility barriers, making political participation more inclusive and fairer for more people.

AI tools are also being used to support voter and civic engagement. Chatbots and conversational AI are showing up in new ways, including early efforts to build AI-powered voter help desks and Q&A platforms.

A campaign chatbot on WhatsApp could answer voters' questions about policies instantly, and with generative AI, the conversation would feel natural and personalized rather than robotic. Though still in the early stages, researchers say advanced AI models could be game-changing by enabling real-time, two-way conversations with voters in multiple languages, making it possible to engage millions of people personally, all at once.<sup>9</sup>

In 2023, a British startup launched an AI-powered “door-knocking” simulation bot to train human canvassers. This allowed volunteers to practice persuasive conversations with a realistic AI voter before knocking on real doors.<sup>10</sup> This kind of training boost could improve the quality of campaign outreach. Soon, campaigns may go further, using AI agents to augment volunteer efforts in contacting voters, but always under human supervision. For example, an AI system could handle routine follow-up messages to remind supporters to vote, freeing human volunteers to focus on high-value personal conversations.

More than improving efficiency, AI can also improve and scale personal interactions without losing the individualized touch that encourages voter connection. Typically constricted by time, human resources, and geography, AI can enable voter outreach to happen more fluidly across platforms and languages and lower barriers such as isolation, scheduling conflicts, or discomfort with face-to-face encounters. Chatbots and simulation bots further allow campaigns to standardize training and messaging, ensuring outreach remains consistent, persuasive, and aligned with strategic goals.

AI is also mobilized by those seeking to protect electoral integrity and empower voters through information. Election watchdog groups and technology companies are using AI more often to find and stop harmful online activity. Social media companies, including Meta, the parent company of Facebook, employed AI systems to track fake accounts, networks of automated bots, and behaviors tied to organized attempts at spreading misinformation. In 2019, Facebook disclosed that it got rid of billions of fake accounts, many of which were caught by automated systems, to prevent manipulation.<sup>11</sup> Such enforcement is far from perfect. Still, without AI, it would have been impossible to police platforms at the scale of billions of users.<sup>12</sup> Likewise, content moderation AI screens for hate speech or incitement in political discourse.

Civil society is also pitching in. Fact-checkers are using AI tools to help debunk false claims faster<sup>13</sup> and other initiatives are leveraging machine learning to flag doctored images or videos<sup>14</sup> before they go viral.

All these applications treat AI as part of the solution,

not merely a source of problems for democracies and elections. When deployed responsibly, AI can strengthen democratic participation by capacitating campaigns to better inform and mobilize the electorate and equipping election integrity defenders with new ways to detect wrongdoing. Solely relying on human monitoring is insufficient in an information ecosystem overwhelmed by the sheer volume and speed of digital content. However, these systems are not perfect, and AI can also be turned into a weapon against fair elections. Therefore, while AI can greatly help protect election integrity, we must carefully examine how it is used and always combine it with human supervision, ongoing improvements to the technology, and strong accountability measures to protect democratic principles.

### New dangers

In spite of its helpful potential, AI also creates serious risks for democracy, especially when it is misused during campaigns and elections. The biggest concern is the creation of highly realistic fake content, like deepfake videos and other AI-generated propaganda, that can fool voters and damage the reliability of information they need to make informed decisions. Deepfakes or AI-generated synthetic content, which can include fake images, audio, and videos, can convincingly replicate the appearance and voices of real individuals. When applied to politics, deepfake technology creates substantial dangers by making it possible to produce fake recordings or staged events intended to harm political competitors, manipulate voter opinions, and diminish faith in credible media and democratic institutions.

Election observers around the world are alarmed at how quickly this threat is growing. The UN has warned that AI-driven disinformation is an escalating threat to democracy,<sup>15</sup> especially as deepfakes have already been used to try to sway elections. Furthermore, when disinformation spreads faster than fact-checks, it creates a first-mover advantage of sorts for bad actors that makes the malicious use of deepfakes more enticing.

For instance, in the U.S., a person who made a deepfake robocall making use of then-President Joe Biden’s voice to discourage voting in key districts during the 2024 elections was subsequently charged with impersonation, as well as election

interference.<sup>16</sup> In India's general elections that same year, deepfake videos falsely showing Bollywood actors Aamir Khan and Ranveer Singh criticizing Prime Minister Narendra Modi went viral before being debunked.<sup>17</sup> In the run-up to Germany's 2024 regional elections, right-wing social media pages pushed AI-generated images and videos of fake "influencers" to quietly seed misinformation into public discourse.<sup>18</sup> Meanwhile, in the Philippines, the so-called "polvoron" deepfake from July 2024 falsely showed President Ferdinand Marcos Jr. using illegal substances,<sup>19</sup> raising concerns that, even though he was not running, the video could hurt support for administration-aligned candidates in the 2025 midterm elections.

Constant exposure to fake or manipulated content can, over time, leave voters cynical or disengaged, especially as it becomes harder for them to tell real from fake. This kind of dissolution of trust makes people less likely to participate in the electoral process. Moreover, when false claims are used to attack political opponents, it does not only hurt individual reputations but also calls into question the fairness of the entire system. Ultimately, if voters cannot trust the basic information they receive during a campaign, such as a video of a candidate speaking, it becomes nearly impossible for them to make an informed choice. An analysis by the Konrad-Adenauer-Stiftung noted that manipulated media "blur reality and erode trust of the public in available information."<sup>20</sup>

In democracies that are well-functioning, using AI-generated fake videos intentionally and with malice, such as fabricated audio clips of a candidate conceding or falsified footage of ballot tampering, threatens an election's legitimacy. These are falsehoods, yes, but they are also targeted efforts to cast a shadow of doubt on the electoral process. In addition, in societies with existing fault lines, more undetectable disinformation can prove especially dangerous. As faith in elections declines, political and even social divides become more pronounced, raising the possibility of unrest, and, in extreme cases, even violence. Furthermore, even when such fakes are eventually debunked, the damage persists.

Compounding the problem is the phenomenon known as the liar's dividend: the idea that the very existence of deepfake technology gives unethical politicians an excuse to dismiss real and

incriminating evidence as "fake." As AI-generated content proliferates, people may become skeptical of the authenticity of genuine audio or video simply because they know convincing fakes are possible.

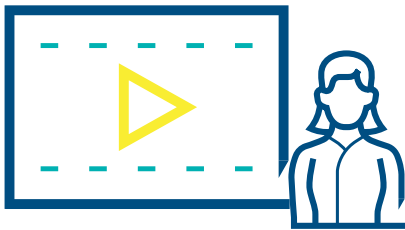
Such a case happened in Mexico. An audio recording leaked on social media exposed a mayor talking about illicit payments for a bid for re-election. Faced with the incriminating audio leak, the mayor declared that it was a deepfake and denied everything. In reality, analysis showed that the audio was 95% likely to be authentic, but just the claim of AI trickery muddied the waters enough to prevent accountability.<sup>21</sup>

As deepfakes become more and more common and AI makes their creation easier and more accessible, we can expect more public figures to cry "fake!" when compromising material emerges. This disintegration of a shared baseline of truth poses a serious challenge for democracies. Without trusting that the evidence is real, even legitimate revelations of wrongdoing may fail to persuade the public or result in justice. In short, AI threatens to deceive people with fake information and cover up real malfeasance by seeding doubt about all information.

Beyond setting off instability and division, the long-term and constant exposure to AI-powered disinformation can lead to normalizing distrust as a baseline political attitude. Rather than judging candidates on their policies, track record, or honesty, voters might become cynical about all political messages. People might begin to perceive elections as simply about those who can spin the best lies instead of those who have the better ideas and resume. When this occurs, people are put off by politics altogether, and that is exactly the environment where demagogues thrive. They capitalize on public disenchantment and political withdrawal, since disengaged populations are more receptive to oversimplified remedies for complex governance issues.

Furthermore, deepfakes are becoming increasingly insidious because they can be tailored for character assassination and incitement. In many countries, female public figures have been targeted by deepfake pornography as a form of harassment: phony explicit videos or images created to humiliate them. For example, several British female politicians became targets of AI-generated

pornographic imagery.<sup>22</sup> According to a 2023 study, 98% of online deepfake videos are pornographic, and the overwhelming majority of over 99%, exploit women and girls.<sup>23</sup> These gendered attacks inflict real personal harm and risk, pushing women out of the public sphere, discouraging their participation not only in elections, but in political life more broadly. Deepfake harassment reveals and aggravates gender inequalities in politics, sending a chilling message to aspiring female leaders that visibility and participation come with significant challenges.



AI-manipulated media could also be used to stoke ethnic or religious tensions, for example, by falsely showing a politician making derogatory remarks about a certain community, potentially inflaming unrest. In the U.S., we have seen how quickly AI-amplified falsehoods can erode public trust along partisan lines. AI-generated fake videos targeting U.S. politicians drew huge attention on social media, with clips of former President Biden and former Vice President Kamala Harris racking up millions of views.<sup>24</sup> During recent natural disasters like hurricanes, AI-generated rumors spread on social networks, sowing distrust in government relief agencies.<sup>25</sup>

In the long term, these examples can impede women and other vulnerable groups from going into public office and divest democracies of diverse leadership. This weakens the legitimacy of political institutions by making them less inclusive and representative, especially since this can adversely impact other vulnerable groups when it comes to religion, gender identity, and race, among others. Just as troubling, when this kind of digital harassment becomes normalized, the public can grow desensitized to it, allowing online abuse, especially against women, to become a routine and unchecked part of political life.

Without intervention, these threats could eclipse AI's benefits for democracy. What is at stake is the integrity and fairness of campaigns and elections, as well as the voters' ability to make informed

choices. That is why there has been a growing push for clearer rules and safeguards around how AI is used in politics.

### Protecting electoral integrity and the case for regulation

As AI's risks have come into sharper focus, the demand for strong electoral regulation has grown among policymakers and civil society, aiming to capture AI's benefits while restraining its worst harms. Few believe that self-policing by political parties or technology companies is enough, given the high stakes for democracy. Governments are amending election and online laws to keep AI from interfering with elections. They are trying to balance encouraging new technology while protecting core democratic principles. The basic idea is simple: elections are the foundation of democracy, so when powerful AI tools are used in this arena, there needs to be oversight to ensure they do not manipulate or distort the actual will of the voters.

Different jurisdictions are pursuing various strategies. The European Union (EU) is the first major power bloc to regulate how AI could threaten elections. Under its Artificial Intelligence Act, any AI system that might influence voting or referendums must meet strict safety rules and face heavy oversight. This indicates that if political campaigns in the EU want to use AI tools, like systems that target specific voters with personalized ads or technology that creates fake videos, they would first have to prove that these tools are transparent, safe, and adequately supervised by actual humans.<sup>26</sup> Providers of such AI systems will probably be required to conduct risk assessments and audits, provide detailed documentation, and ensure a level of explainability so that the technology's impact on voters can be understood and scrutinized. The EU is moving toward a regime where manipulative AI in politics is not outright banned but is tightly supervised under regulatory watch.

The EU has also established additional electoral safeguards that go beyond the AI Act, featuring new transparency obligations for political campaign advertisements. In early 2024, the EU passed the Regulation on the Transparency and Targeting of Political Advertising to clarify how political actors use online advertisements and customized messaging. Under these rules, all political ads must be clearly labeled as political and include



Photo taken during the breakout session on "Harnessing AI for Good Governance" during the 2024 KAS Solutions Conference.

information on who paid for them, to which election or issue they relate, and whether any targeting was used.<sup>27</sup> Advertisers also need to provide transparency notices with details like how much they spent and who funded them. The regulation also cracks down on microtargeting. Political ads can only target people online if they use data directly from those individuals, and only if those people have specifically agreed to be targeted for political content. Moreover, sensitive personal information such as ethnicity, religious beliefs, or political opinions is entirely prohibited from being used for targeting purposes.<sup>28</sup> The EU also plans to maintain a public online repository of all political advertisements and ban political advertising from non-EU sources in the final weeks leading up to elections.<sup>29</sup> These measures seek to prevent the kind of opaque, invasive profiling that the Cambridge Analytica scandal<sup>30</sup> exemplified and to give voters more agency, or at least awareness, concerning the customized messages they receive.

Instead of simply banning AI, the EU applies a more measured approach. The regional bloc regulates how AI is used in elections, concentrating on maintaining transparency and proper monitoring. By having clear guidelines for when and how AI can be put to work, the EU shows that technology does not have to be a threat to democracy. It can actually help strengthen it when used responsibly. This puts voters in the forefront by giving them real insight into how their data is used, preventing companies from taking advantage of their personal data, and keeping foreign actors from interfering with elections. It is a way to shield democracy from tech-related risks without undermining people's trust or letting politicians get away with dishonest messaging. For other democracies wrestling with these same issues, the EU's model provides a roadmap.

In the Philippines, a country that has, and is, intensely experiencing fake news and disinformation in past elections, authorities moved to rein in AI-driven abuse before the 2025 midterm polls. The country's Commission on Elections (COMELEC) issued new guidelines (Resolution No. 11064) in late 2024 that impose strict rules on using social media and AI for campaigns. These rules break ground in several areas.

For one, all political candidates and parties must register their official campaign pages and social media accounts with the election commission to create an authenticated roster of online campaign channels. This makes it easier for regulators to monitor propaganda and for the public to identify official content.

Second, the COMELEC mandates that any AI-generated campaign content carries a clear on-screen disclaimer on the content itself, stating that it has been manipulated by AI. The rule even requires disclosing how the content was manipulated and confirming that any real person depicted gave their consent. In practice, if a politician wanted to use an AI-generated image of their opponent or of a celebrity endorser, they would have to label it visibly.

Also, Philippine rules do not allow bots or fake accounts for online campaigning. Campaign teams need to check their content to make sure that they are not accidentally spreading false information through AI-generated fake profiles or automated posts. Social media platforms are expected to work with the COMELEC to take down campaign pages that are not properly registered or do not follow the rules. Thus, the COMELEC recommends digital watermarking or authenticity verification on campaign images and videos.<sup>31</sup>

These provisions from the Philippines collectively represent one of the more comprehensive efforts to directly regulate AI in political campaigns. The aim is to catch violators after the fact and deter malicious uses of AI in the first place by setting clear legal boundaries. Candidates can no longer flood the public with fake content without facing consequences. These rules will require vigorous enforcement and close cooperation from tech companies. Still, they recognize an important reality that extraordinary technological threats call for proactive governance.

Other democracies are moving in the same direction, with growing support for labeling requirements on AI-generated political content. Even some U.S. states, like California, have enacted laws making it illegal to distribute malicious deepfakes about a candidate near an election.<sup>32</sup> There are also appeals for campaign ethics pledges where parties voluntarily commit not to use deepfakes or weaponize AI against opponents.<sup>33</sup> Within the international community, groups like the United Nations Development Programme and global election observers are working on guidelines and toolkits to help handle AI's effects on elections.<sup>34</sup>

These efforts are forming an emerging consensus that democratic norms and laws must be updated for the AI era. Free-speech absolutism has its limits when highly sophisticated fakery and micro-manipulation are in play. Some guardrails are necessary to preserve a meaningful ecosystem of public discourse in elections.

## Toward a democratic digital future

Based on the examples and analyses above, it is clear that making the most of AI's potential for democracy, while minimizing the risks, calls for a multi-pronged approach. The recommendations below intend to offer a path forward for policymakers, election officials, platforms, and civil society.

### 1 Strengthen and update legal frameworks.

Governments need to set clear laws to stop the harmful use of AI in elections, especially when it comes to deepfakes and other deceptive content. These laws should also require full disclosure when political messages are created, altered, or enhanced with AI. Equally important is accountability. Candidates and campaign teams that knowingly spread AI-generated disinformation or deploy bots to sway public opinion should face serious penalties, including fines, disqualification, or in extreme cases, criminal charges. Rules alone cannot eliminate every misuse, but they can deter most bad actors and equip authorities to act when boundaries are crossed.

### 2 Ensure solid enforcement and oversight.

Rules do not mean much without enforcement. Election bodies, regulators, and civil society need

the tools and skills to monitor digital campaigns in real time. That means setting up dedicated AI teams or command centers to spot deepfakes, bot networks, and harmful content and take them down fast. Close coordination with tech platforms is essential, as having direct lines of communication helps speed up the removal of fake accounts and the flagging of misleading posts. Election officials should also have the power to audit how campaigns use data and AI, and require proof that these tools meet established standards when concerns arise. International observer teams should add digital forensics experts to monitor online behavior. Oversight has to evolve with AI, calling for fresh skills and agile responses. To avoid censorship and bias, takedowns should follow clear, public rules with independent oversight and an appeals process. Transparency reports and multi-stakeholder reviews can help ensure enforcement stays accountable and respects free speech.

### 3 Promote transparency and empower voters.

Governments and platforms should make it easier for the public to see how AI shapes the political content they encounter. Social media companies, for example, should maintain public ad libraries that show all political ads, who they targeted, and whether AI was used in creating them. Deepfakes and fake accounts flagged for misinformation should also be disclosed (e.g., "This content was flagged as AI-generated and removed for misinformation"). Governments can help by running public service campaigns that explain how AI-generated fakes work and how to spot them. Media and AI literacy should be part of civic education, so people learn early how to assess what they see online. In some countries, volunteers and journalists already do real-time fact-checking during debates and viral events. These efforts should be expanded. The goal is to build public awareness and give people tools, like browser plugins or verification apps, to verify content independently. A well-informed, skeptical public is the strongest defense against digital manipulation.

### 4 Encourage ethical innovation and best practices.

Not all solutions must be top-down. Political parties and campaigns can collectively adopt ethical codes of conduct for AI use. For example, parties might agree that they will not create deepfakes of

opponents, will not knowingly distort information even when using AI, and will promptly correct or rescind any misinformation their AI tools create. Pledges could be made public so that voters and the media can hold campaigns to account. Likewise, AI developers and technology firms should integrate ethical safeguards, including watermarking and detection features, to make AI outputs easier to differentiate from human content. Increased investment in advanced deepfake detection technology by governments, industry, and philanthropy is crucial, ensuring continuous progress in the ability to detect fakes as they emerge. In the same way, stakeholders should explore ways to use AI positively: for instance, social media platforms could deploy auditors to identify algorithmic bias during elections, to ensure their recommendation systems are not disproportionately amplifying extreme or false content. By sharing best practices internationally through conferences, multi-stakeholder forums, and UN guidelines, democracies can learn from one another what works to mitigate AI threats, while still leveraging AI for civic good.

**5 Advance international cooperation and norms.**

The global flow of digital information means no country can face these challenges alone. Foreign

interference in elections, often amplified by AI, is a serious and ongoing threat, as seen in Russian campaigns in the West and reported Chinese influence in parts of Asia. Democracies should work together to set clear norms, or even formal treaties, against the malicious use of AI in politics. Just as cyberattacks on critical infrastructure are broadly condemned, we need agreements that, for example, treat the spread of deepfakes to disrupt elections as unacceptable behavior. International organizations, such as the ASEAN, EU, and the UN, can give support by helping joint research on how AI affects democracy and by coordinating election support, whether that means sending technical experts or providing monitoring tools. Defending democracy in the age of AI will take constant vigilance, cooperation, and the ability to adapt.

AI is often characterized as a double-edged sword. In campaigns and elections, that comparison is on point. It has the power to encourage political participation, but it can just as quickly distort truth and break trust. Whether AI strengthens democracy or undermines it will depend on the choices made by policymakers, technologists, campaigners, and citizens in the years to come. With smart regulations, transparency, and a shared sense of responsibility in our arsenal, we can push AI to aid voters and the public rather than those who seek to manipulate or deceive.

---

## References

- 1 United Nations Development Programme, "A 'Super Year' for Elections," UNDP Feature, January 2024, <https://www.undp.org/super-year-elections>.
- 2 Microtargeting is the practice of using AI and data analytics to categorize voters into ultra-specific subgroups and convey tailored messages according to their distinct characteristics or behaviors.
- 3 Maria Curi, "Exclusive: AI turbocharges campaign fundraising," *Axios*, January 30, 2024, <https://www.axios.com/2024/01/30/ai-campaign-fundraising-democrats-chatgpt>.
- 4 Ibid.
- 5 Kenneth Christiane L. Basilio, "PHL midterm elections at risk from deepfakes, political analysts say," *BusinessWorld Online*, January 28, 2025, <https://www.bworldonline.com/the-nation/2025/01/28/649536/phl-midterm-elections-at-risk-from-deepfakes-political-analysts-say>.
- 6 Ibid.
- 7 Florian Foos, "The Use of AI by Election Campaigns," *LSE Public Policy Review* 3, no. 3 (2024): 8, <https://doi.org/10.31389/lseppr.112>.
- 8 Mariyam Alavi, "BJP Shared Deepfake Video On WhatsApp During Delhi Campaign," *NDTV*, February 20, 2020, <https://www.ndtv.com/india-news/in-bjps-deepfake-video-shared-on-whatsapp-manoj-tiwari-speaks-in-2-languages-2182923>.
- 9 Foos, "The Use of AI by Election Campaigns."
- 10 Ibid.
- 11 Jeremy Kahn, "Meet the A.I. that helped Facebook remove billions of fake accounts," *Fortune*, March 4, 2020, <https://fortune.com/2020/03/04/facebook-a-i-fake-accounts-disinformation>.
- 12 Rem Darbinyan, "The Growing Role Of AI In Content Moderation," *Forbes*, June 14, 2022, <https://www.forbes.com/councils/forbestechcouncil/2022/06/14/the-growing-role-of-ai-in-content-moderation>.

- 13 Lydia Morrish, "Fact-Checkers Are Scrambling to Fight Disinformation With AI," *WIRED*, February 1, 2023, <https://www.wired.com/story/fact-checkers-ai-chatgpt-misinformation>.
- 14 "Here's how we're using AI to help detect misinformation," Meta AI, November 19, 2020, <https://ai.meta.com/blog/heres-how-were-using-ai-to-help-detect-misinformation>.
- 15 Basilio, "PHL Midterm Elections at Risk."
- 16 David Shepardson, "US political consultant indicted over AI-generated Biden robocalls," *Reuters*, May 24, 2024, <https://www.reuters.com/world/us/us-political-consultant-indicted-over-ai-generated-biden-robocalls-2024-05-23>.
- 17 Reuters, "Aamir Khan and Ranveer Singh criticising PM Modi! Think again, that's deepfake drama," *The Economic Times*, April 22, 2024, <https://economictimes.indiatimes.com/news/elections/lok-sabha/india/deepfakes-of-bollywood-stars-spark-worries-of-ai-meddling-in-india-election/articleshow/109487075.cms>.
- 18 Kathrin Wesolowski and Joscha Weber, "Fact check: AI influencers targeting German elections," *DW News*, February 5, 2025, video, 8:52, <https://www.dw.com/en/fact-check-ai-influencers-targeting-german-elections/video-71514325>.
- 19 Alexis Romero, "AI experts debunk Marcos' alleged drug snorting video," *The Philippine Star*, September 22, 2024, <https://www.philstar.com/headlines/2024/09/22/2387005/ai-experts-debunk-marcos-alleged-drug-snorting-video>.
- 20 Ferdinand Alexander Gehringer, Christopher Nehring, and Mateusz Łabuz, "The Influence of Deep Fakes on Elections: Legitimate Concern or Mere Alarmism?" *Konrad-Adenauer-Stiftung Monitor*, May 10, 2024, <https://www.kas.de/en/monitor/detail/-/content/the-influence-of-deep-fakes-on-elections>.
- 21 "The Mexican Mayor and a Deepfake Scandal," *BBC World Service*, March 14, 2024, 17:30, <https://www.youtube.com/watch?v=i0wwR6kxYQY>.
- 22 Jim Waterson, "British female politicians targeted by fake pornography," *The Guardian*, July 1, 2024, <https://www.theguardian.com/technology/article/2024/jul/01/british-female-politicians-targeted-by-fake-pornography>.
- 23 Leslie Graves, "The Pathfinder: Deepfakes," *PBS*, September 12, 2024, <https://www.pbs.org/wnet/preserving-democracy/2024/09/12/deepfakes>.
- 24 Sam Stockwell, Megan Hughes, Phil Swatton, Albert Zhang, Jonathan Hall, and Kieran, "AI-Enabled Influence Operations: Safeguarding Future Elections," *Centre for Emerging Technology and Security*, November 2024, <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-safeguarding-future-elections>.
- 25 Esther Chan, Hardik Vyas, and Lydia Morrish, "Hurricane Milton: Fact-checking conspiracies, rumors and AI images," *Reuters*, October 12, 2024, <https://www.reuters.com/fact-check/hurricane-milton-fact-checking-conspiracies-rumors-ai-images-2024-10-11>.
- 26 "Artificial Intelligence Act: MEPs adopt landmark law," *News | European Parliament*, March 13, 2024, <https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>.
- 27 "Regulation (EU) 2024/900 of 13 March 2024 on the Transparency and Targeting of Political Advertising," March 18, 2024, Arts. 7 and 12, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R0900>.
- 28 Ibid.
- 29 "Parliament adopts new transparency rules for political advertising," *News | European Parliament*, February 27, 2024, <https://www.europarl.europa.eu/news/en/press-room/20240223IPR18071/parliament-adopts-new-transparency-rules-for-political-advertising>.
- 30 Joe Westby, "The Great Hack: Cambridge Analytica is Just the tip of the iceberg," *Amnesty International*, July 24, 2019, <https://www.amnesty.org/en/latest/news/2019/07/the-great-hack-facebook-cambridge-analytica>.
- 31 "Resolution No. 11064: Guidelines on the Use of Social Media, Artificial Intelligence, and Internet Technology for Digital Election Campaigns, and the Prohibition and Punishment of Its Misuse for Disinformation and Misinformation, in Connection with the 2025 National and Local Elections and the BARMM Parliamentary Elections," *Commission on Elections*, September 25, 2024, [https://comelec.gov.ph/php-tpls-attachments/2025NLE/Resolutions/com\\_res\\_11064.pdf](https://comelec.gov.ph/php-tpls-attachments/2025NLE/Resolutions/com_res_11064.pdf).
- 32 "AB 2655: Defending Democracy from Deepfake Deception Act of 2024," *CalMatters*, 2023–2024 Regular Session, chaptered September 17, 2024, [https://calmatters.digitaldemocracy.org/bills/ca\\_202320240ab2655](https://calmatters.digitaldemocracy.org/bills/ca_202320240ab2655).
- 33 Clothilde Goujard, "EU political parties promise to steer clear of deepfakes ahead of election," *Politico*, April 10, 2024, <https://www.politico.eu/article/eu-political-parties-promise-to-steer-clear-of-deepfakes-ahead-of-election>.
- 34 "Defending Information Integrity: Actions for Election Stakeholders," *UNDP*, March 2024, [https://www.undp.org/sites/g/files/zskgke326/files/2024-03/24119\\_undp\\_actioncoalition\\_v06\\_rc\\_hq-spreads.pdf](https://www.undp.org/sites/g/files/zskgke326/files/2024-03/24119_undp_actioncoalition_v06_rc_hq-spreads.pdf).



*Jaemark Tordecilla during the 2024 KAS Solutions Conference.*

# Building AI Tools for Investigative Journalism and Public Accountability

**Jaemark Tordecilla**

---

AI is no longer a future consideration for journalism. In the Philippines, it is already being used in the field to address practical challenges of public accountability and reporting. Over the past year, I have worked on a series of AI tools developed specifically for use in the local media and civic data ecosystem. These tools were developed in response to specific problems encountered by journalists and researchers. Audit reports are often lengthy and difficult to scan for anomalies, so we designed a custom GPT to make these documents easier to analyze. Budget documents are massive and poorly structured for analysis by non-specialists, prompting the development of a bot that allows users to query the national budget and break down its complexities. Investigative reports typically reach only a limited segment of the population unless adapted into video or audio, so we built a workflow to convert long-form investigative stories into multimedia formats for broader public reach.

Although international tools and platforms already exist to assist with some of these tasks, these tools are rarely optimized for local formats or use cases. As a result, reporters, editors, and civil society partners are left to adapt general-purpose tools for their own needs, often with limited success.

This work was driven not only by a desire to share practical tools and workflows but also to highlight that innovation in journalism, especially using AI, can come from anywhere, even in resource-constrained and challenging settings. Newsrooms in the Global South are not simply passive users of these technologies. In fact, the AI tools featured in this article were created despite constraints—limited computational resources, a lack of specialized staff, and complex local data—showing that innovation can thrive even under difficult

conditions and lead to more focused, user-centered, and efficient solutions.

At the same time, the risks posed by AI, including disinformation, regulatory overreach, and editorial compromise, are magnified in environments with weak institutions and shrinking press freedom. Building AI tools is only part of the work. Ensuring that these serve the public interest requires attention to governance, inclusion, and long-term sustainability.

In this article, I present three case studies that demonstrate how localized AI tools can help address the challenges of public accountability reporting. I also consider the policy implications of this work and offer recommendations for how we can navigate the next phase of the AI revolution in journalism.

### **AI and public accountability: the COA Beat Assistant**

Government audit reports are among the most important documents for public accountability. In the Philippines, the Commission on Audit (COA) releases thousands of pages of audit observations every year, one for each agency or government unit. However, these reports are often difficult to navigate. These documents follow no standard template, vary widely in length and structure, and contain dense financial information that is not always easy to interpret. Still, journalists and civil society groups rely on COA reports as a key resource for identifying irregularities or red flags in government spending. But due to resource limitations and time constraints, many reports go unread or are only lightly reviewed.

To address this, I developed the COA Beat Assistant, a custom tool powered by OpenAI's GPT models via ChatGPT. It is free to use, publicly accessible, and specifically designed to extract summaries of significant audit observations from COA executive summaries.<sup>1</sup> The tool was configured with strict parameters to prioritize user control and data security. It processes only the content of the uploaded document and does not access or reference any external sources. No data is stored after a session ends, and because the tool runs on OpenAI's ChatGPT, it inherits the platform's built-in safeguards against data retention and misuse. Furthermore, its utility depends entirely on the

document that the user uploads—it does not fetch data on its own. It is up to users to ensure they are uploading official, legitimate COA reports, which are publicly available on the Commission's website. While the tool can technically be cloned or modified, its public-interest design, narrow functionality, and lack of persistent data storage limit its potential for misuse.

When a document is uploaded, the assistant first asks the user if they want to generate a summary of the audit observations. If the user agrees, the tool scans the "Summary of Significant Audit Observations and Recommendations" section and produces a list of one-sentence summaries, each tagged with its corresponding observation number. The user can then request details for any of the individual observations.

The tool proved useful in real-world tests. For example, Llanesca Panti of GMA News Online used the assistant to review audit reports while monitoring a congressional hearing. She was able to flag entries that would later guide her reporting.<sup>2</sup> In another instance, Columbia University investigative journalism professor Sheila Coronel used the tool to find audit anomalies in local governments in Mindanao, including questionable Gender and Development fund spending.<sup>3</sup> When applied to the 2021 COA report on the Department of Education (DepEd), it surfaced the observation on overpriced laptops, an issue that became a national controversy after it was reported by the media.

While the COA Beat Assistant helped journalists quickly locate leads and identify promising story elements, its limitations reinforced the importance of editorial oversight. The Assistant's summaries were only as accurate as the formatting of the original document allowed. Some observation numbers were dropped or misaligned, and on occasion, figures were extracted from incorrect columns. Journalists testing the tool were quick to verify these summaries against the original documents, using the assistant not as a final authority, but as a guide to help them decide where to look.

These scenarios reinforce the value of designing tools around how experienced and credible journalists actually work. The Assistant was built not just to generate summaries, but to support a

workflow where journalists can scan key observations, follow leads into full reports, and validate findings with the source document. The COA Beat Assistant did not replace editorial judgment; it created new entry points for it. In doing so, it saved time without compromising rigor. Designing tools in this way, with the end-user's process in mind, is critical to ensuring that AI enhances rather than undermines journalistic standards.

The COA Beat Assistant was released in January 2024 in a blog run by the Computational Journalism Lab at Northwestern University.<sup>4</sup> It was later picked up by the Reuters Institute for the Study of Journalism at the University of Oxford.<sup>5</sup> Since then, the tool has been far-reaching and influential. Zach Seward, the AI newsroom leader of the New York Times, cited it as one of the examples that “give us a sense of the technology's greatest promise for journalism” in presentations at events like SXSW<sup>6</sup> and the International Symposium on Online Journalism.<sup>7</sup> The Assistant was also highlighted in the Reuters Institute Digital News Report 2024 country report for the Philippines,<sup>8</sup> and was included among the “Top Tools and Case Studies to Get You Started” on generative AI in WAN-IFRA's Innovation in News Media World Report 2024-25.<sup>9</sup> It will be featured in the upcoming book “The Science of First: AI's Promise to Forecast the News” by Francesco Marconi and Erin Riglin. After making the tool publicly available, I was also asked to lecture at the inaugural AI Journalism Lab at the Craig Newmark Graduate School of Journalism at The City University of New York (CUNY), teaching newsroom innovation experts about my work.

There are two things that resonated with people about my piece. First, by documenting the journey in developing the tool in intricate detail, I was able to show that it was possible to design and build one's own AI solutions even with limited resources. While it discussed building a specific tool for a particular use case in Philippine newsrooms, it was this level of specificity that actually made the experience quite universal, because as I wrote in the article for the Reuters Institute, similar tools “could be created for all sorts of corpora of documents to make life easier for investigative journalists.”<sup>10</sup>

The second reason is more implicit. It involved recognizing what I was good at (journalism) and

what AI was good at (processing information), then figuring out how to integrate these strengths into an effective workflow. This approach provided reassurance to journalists hesitant about using AI in the newsroom as well as a blueprint on how to design and build these tools for their own context.

These experiences with the COA Beat Assistant reflect broader concerns that media organizations in the Global South must confront. As AI becomes embedded in more newsroom workflows, it is vital to ensure that these technologies do not exacerbate existing inequalities or undermine democratic accountability. Resource-constrained newsrooms often face barriers in accessing AI tools; yet, the need for oversight and transparency in governance is more urgent than ever. Tools like the COA Beat Assistant show that it is possible to develop effective AI applications even in challenging conditions, but sustaining this work will require stronger institutional support, equitable access to infrastructure, and international frameworks that promote both innovation and press freedom.



### AI and fiscal democracy: the Budget Bot

The Philippines is widely recognized for having one of the most transparent budget processes in the world.<sup>11</sup> According to the 2023 Open Budget Survey, the Philippines is among the top countries in terms of budget oversight and public access to budget documents.<sup>12</sup> However, high transparency scores do not automatically translate to broad participation or accessibility.

The national budget files, while publicly available, remain difficult to use for most citizens and even for many journalists. The Philippine national budget is released as a downloadable spreadsheet by the Department of Budget and Management (DBM). However, the format is far from user-friendly. Each year's document is made up of hundreds of thousands of line items and dozens of columns, with varying data conventions across agencies. The

files are too large to open on most computers and require either business intelligence software or proficiency in data science tools like Python or R to navigate. In addition, despite high international rankings for budget transparency, public participation in fiscal oversight remains limited. Journalists, civil society organizations, and even some government agencies often rely on pre-summarized versions of the budget or tips from insiders to identify anomalies or trends. Thus, the budget is rarely examined in full, and comprehensive analysis is typically confined to a small circle of specialists.

The Budget Bot was created to improve and change this situation.<sup>13</sup> It enables users to query the national budget using plain language. For example, a user could ask, “What’s the budget for the University of the Philippines?” and the bot would generate the necessary code to find and display the relevant line items. It draws from cleaned-up CSV versions of the official spreadsheet of the budget and includes a simplified knowledge base that helps it match informal queries to official agency names. The tool is publicly available and free to use, powered by OpenAI’s ChatGPT with Code Interpreter. Its configuration files and cleaned datasets are also published on GitHub, allowing journalists, researchers, and civic groups to replicate, adapt, or build on the tool for their own use.

Getting the tool to work reliably took several adjustments. At first, the original Excel budget files were simply too large to handle easily. Converting them into CSV (Comma-Separated Values) format significantly improved performance and cleaning up extra columns, such as those containing reference codes and metadata, made the searches faster. Another key change was creating a reference list of government agencies, which helped the bot better match user queries with the correct entries. These steps made the tool more accurate and easier to use.

The design of the Budget Bot prioritized usability. It was not enough to answer questions correctly; the tool had to explain which dataset it was using, what search terms were applied, and how the results were calculated. Transparency was a core principle. If the user requested budget figures for flood control, the bot would display not just totals, but also the matching line items and their descriptions.

This allows users to double-check the output and incorporate it into their reporting with confidence.

Budget Bot is not a replacement for expert budget analysis, but it enables broader access to budget data. It has already been used to identify confidential and intelligence fund allocations, compare regional education budgets, track infrastructure allocations in specific provinces, and surface discrepancies between proposed and approved funding across fiscal years. The Budget Bot Graph Gallery on Instagram presents data visualizations illustrating these, as well as showcasing important findings using data from the tool, such as infrastructure allocations, shifts in advertising expenses by government agencies, and agency-level breakdowns of confidential fund allocations.<sup>14</sup> These use cases highlight its potential as a practical tool for both journalists and researchers.

To support further use, I published the cleaned datasets and the bot’s configuration files in a public GitHub repository.<sup>15</sup> This transparency allows other journalists, researchers, and civic technologists to verify the methodology, adapt the tool to their own contexts, and conduct deeper or comparative analyses using consistent, reproducible data.

The Budget Bot exemplifies a broader shift in how people interact with public data with the help of AI. Now that it is possible to interact conversationally with the national budget through AI, a natural follow-up question arises: can we do the same for city budgets, procurement records, or infrastructure plans? Furthermore, why are government budget datasets not available in accessible and user-friendly formats? AI tools like Budget Bot challenge existing assumptions and encourage citizens to demand more transparent, usable, and responsive public information systems. This is a foundational use case not only for the Philippines but also for newsrooms and civic actors globally. It demonstrates how large language models can be locally adapted to interface with public datasets, even in low-resource settings. The Budget Bot shows that it is possible to reduce the complexity of massive government documents and make them searchable through simple, guided conversation.

The learnings from this work have been recognized by the international media community. For

newsrooms around the world, it provides a blueprint for developing AI tools that are grounded in journalistic practice, designed for real users, and capable of expanding access to essential public information.

At a broader level, Budget Bot demonstrates how AI tools can be a shared civic resource: openly published, easily reused, and adaptable for a range of users beyond its original newsroom context. It embodies a "build once, run many times" approach, where the value increases the more people use, extend, or replicate the tool.

For this kind of work to grow, it needs more than individual effort. It requires ecosystems that support collaboration, infrastructure, and long-term sustainability so that tools built in one newsroom can benefit many others. These same ecosystems are necessary to ensure that tools like Budget Bot do not remain isolated experiments but become part of a broader shift in how AI can be used to support democracy. Sustaining and scaling these efforts demands investment, governance, and inclusive frameworks that address the specific needs of media organizations in the Global South.

### **AI and multimedia storytelling: "The Making of Edgar Matobato" Project**

In addition to data-driven tools, I explored the use of generative AI to repurpose investigative journalism for wider distribution. One case study was "The Making of Edgar Matobato," a long-form narrative written by Sheila Coronel and published by the Philippine Center for Investigative Journalism (PCIJ), which recounts the story of a self-confessed hitman who alleged he was part of a death squad linked to then-Davao City Mayor Rodrigo Duterte during the country's drug war. The piece was powerful and well-researched, but like many long-form investigations, its reach was limited to a niche audience already familiar with PCIJ's work.

With permission from the author, I adapted the story into a video format using a suite of generative AI tools.<sup>16</sup> AI voice synthesis software was used to narrate the story with Filipino-accented voices. Images were generated in a consistent visual style using AI image generators and then animated to match the tone and structure of the narrative.<sup>17</sup> The result was a 30-minute video produced in under two days. By comparison, a traditional production

workflow involving voice talent, illustrators, and editors would have taken weeks and required far more resources.

The project highlighted a set of new opportunities. First, it enabled the same story to reach audiences on video-first platforms where text-based journalism tends to be ignored. Second, it opened the possibility for small newsrooms to experiment with multimedia storytelling without having to invest in full-scale production teams. The project also prompted a compelling discussion among journalists and educators. While some were impressed with the synthetic voices and consistency of the visual style, others pointed to concerns about monotony, mispronunciations, and the ethics of AI-generated visuals.

These concerns are important. Generative AI is not a substitute for editorial or artistic judgment. One of the main challenges in the project was ensuring that the AI-generated content accurately reflected the tone and factual detail of the original story. This required manually correcting voiceovers, careful editing of image prompts, and post-processing of animation sequences. Editorial control remained essential throughout the process.

The project also sparked a broader conversation about the relationship between AI and human creativity. Many colleagues noted that while AI could speed up production and expand reach, the emotional nuance, ethical judgment, and storytelling discipline that define good journalism still depended entirely on human decisions. Rather than replacing journalists, the tools offered an opportunity to reimagine workflows where AI handles repetitive production tasks. While this creates efficiencies, it also raises important questions about labor, authorship, and the role of creative professionals in an AI-augmented newsroom. Editorial responsibility, however, remains firmly with humans, especially in maintaining narrative integrity and audience trust.

It also raised questions about labeling and disclosure. Should audiences be told which parts of a story were AI-generated or manipulated? How do we ensure that synthetic content does not distort the meaning of real events? Moreover, how do we balance the efficiency gains of AI with the responsibility to credit human artists—especially when generative models are trained on vast

datasets of existing creative work to emulate the art, images, voices, and styles of real people, often without the artists' knowledge or consent? Again, this is a foundational use case for journalism worldwide. The workflows and lessons from this project have resonated beyond the Philippines and are cited as a model for integrating generative tools in ethically-grounded editorial work.<sup>18</sup>

For policy and newsroom practice, this case shows the importance of developing clear standards around the use of AI in storytelling. This includes labeling requirements, ethical review processes, and transparency protocols that allow audiences to understand what they are consuming. At the same time, it opens the door for more inclusive journalism. By lowering production barriers, AI can make it easier to tell important stories to audiences who might otherwise never encounter them. While this project focused on adapting a single long-form story, the workflow can be applied across various formats and topics, expanding the reach of public-interest reporting. As with the other tools in this article, the Matobato project is a reminder that inclusive, values-driven AI journalism must be built intentionally, and with an eye toward the broader governance frameworks that shape media freedom. These experiments also show that journalists and newsrooms in the Global South can play a leading role in shaping the norms, practices, and expectations for how AI is used in public-interest media worldwide.



## Conclusion

For the past few years, news organizations in the Global South have confronted a unique set of challenges in their role as democratic watchdogs. These include political pressure from populist and authoritarian rulers, the dominance of tech platforms over online audiences and advertising revenues, a fragmented media landscape, and shrinking newsroom resources. Amid this increasingly constrained environment, AI is beginning to reshape the information ecosystem

globally, offering both critical opportunities and serious threats.

Much of the global discourse on AI and journalism centers on the risks: disinformation, deepfakes, and erosion of editorial integrity. While these are valid concerns, they are compounded in low-resource settings by deeper vulnerabilities: lack of technical infrastructure, affordability barriers, and growing legal threats that could be weaponized through overreaching AI regulation.

The tools presented in this article address some of these issues directly. Each project was designed to work within tight constraints and to complement human editorial oversight. Together, they provide early evidence that responsible AI use in journalism is possible even without large budgets or engineering teams. They demonstrate how resource-constrained newsrooms can design tools and workflows that directly improve public-interest reporting, contributing to faster identification of story leads, greater accessibility to complex government datasets, and more cost-effective production of multiplatform content. By automating time-consuming tasks, these tools could enable journalists to produce more stories with fewer resources. They also make it easier to create formats that resonate with different audiences, such as videos, visuals, or interactive summaries, leading to higher levels of engagement and broader reach.

At the same time, these tools lower the technical barriers that typically limit data-driven reporting to large, well-funded newsrooms. Reporters in smaller, regional, or under-resourced outlets gain new capacity as the tools can democratize both information access and technological capability, expanding who gets to do journalism, what stories get told, and how the public is able to engage with them. More importantly, these use cases suggest that the Global South can lead and not just follow in defining how AI is applied to support democracy, transparency, and civic engagement. By grounding innovation in real-world newsroom needs, these efforts show how AI can strengthen, not replace, the core values of journalism and serve as a force multiplier for accountability and public service.

Still, sustaining and expanding this work to integrate AI into journalism requires more than prototypes. It demands collaborative infrastructures, supportive

policy environments, and inclusive governance frameworks that recognize the needs and contributions of journalists working in politically and economically precarious contexts. If properly

supported, these tools can help build a more equitable future for journalism: one in which AI serves, rather than threatens, public accountability, democratic values, and the role of the free press.

---

## References

- 1 Access the COA Beat Assistant at <https://chatgpt.com/g/g-6UIXp281z-coa-beat-assistant>.
- 2 Jaemark Tordecilla, "I created an AI tool to help investigative journalists find stories in audit reports. Here's how I did it," *Reuters Institute for the Study of Journalism-University of Oxford*, February 16, 2024, <https://reutersinstitute.politics.ox.ac.uk/news/i-created-ai-tool-help-investigative-journalists-find-stories-audit-reports-heres-how-i-did-it>.
- 3 Ibid.
- 4 Jaemark Tordecilla, "Teaching a Custom GPT to Read Audit Reports and Support Watchdog Journalism," *Generative AI in the Newsroom via Medium*, January 30, 2024, <https://generative-ai-newsroom.com/teaching-a-custom-gpt-to-read-audit-reports-and-support-watchdog-journalism-1e124196ca15>.
- 5 Tordecilla, "I created an AI tool."
- 6 Zach Seward, "AI news that's fit to print: My talk at SXSW 2024," *ZMS*, March 11, 2024 <https://www.zachseward.com/ai-news-thats-fit-to-print-sxsw-2024/>.
- 7 Zach Seward, "Creating structure with generative AI: My talk at ISOJ 2024," *ZMS*, April 12, 2024, <https://www.zachseward.com/creating-structure-with-generative-ai/>.
- 8 Yvonne T. Chua, "Digital News Report 2024: Philippines," *Reuters Institute for the Study of Journalism-University of Oxford*, June 17, 2024, <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2024/philippines>.
- 9 "Innovation in News Media World Report 2024-25," *WAN-IFRA*, May 29, 2024, <https://wan-ifra.org/insight/innovation-in-news-media-world-report-2024-25/>.
- 10 Tordecilla, "I created an AI tool."
- 11 Jean Mangaluz, "Philippines ranks first in Asia for budget transparency," *Philippine Daily Inquirer*, May 30, 2024, <https://business.inquirer.net/461567/philippines-ranks-first-in-asia-for-budget-transparency>.
- 12 "Open Budget Survey 2023 Country Results: Philippines," <https://internationalbudget.org/open-budget-survey/country-results/2023/philippines>.
- 13 Access Budget Bot at <https://chatgpt.com/g/g-JbtdqTBNp-budget-bot>.
- 14 View @budgetbotph, "Budget Bot Graph Gallery," on Instagram at <https://www.instagram.com/budgetbotph/>.
- 15 Jaemark Tordecilla, J, "Budget Bot [Code repository]," GitHub, <https://github.com/tordecilla/budget-bot>.
- 16 Jaemark Tordecilla, "The Making of Edgar Matobato: PCIJ experiments with using AI for storytelling," *PCIJ*, January 30, 2025, <https://pcij.org/2025/01/30/the-making-of-edgar-matobato-pcij-experiments-with-using-ai-for-storytelling/>.
- 17 Jaemark Tordecilla, "How I used generative AI to turn a powerful long-form story into an illustrated video," *Reuters Institute for the Study of Journalism-University of Oxford*, February 7, 2025, <https://reutersinstitute.politics.ox.ac.uk/news/how-i-used-generative-ai-turn-powerful-long-form-story-illustrated-video>.
- 18 Ezra Eeman, "Wayfinder Media Trends Report: navigating tomorrow," *International Journalism Festival*, April 10, 2025, accessed at <https://www.journalismfestival.com/programme/2025/wayfinder-trends-report-navigating-tomorrow>.



*Dominic Vincent Ligot during the 2024 KAS Solutions Conference.*

# The Case for Sovereign AI: A Philippine Perspective

**Carljoe Javier and Dominic Vincent Ligot**

---

Kai-Fu Lee, a Taiwanese businessman and computer scientist, named China and Silicon Valley in the U.S. as the AI superpowers. The title of his book is prescient and draws appropriate attention to the fact that both China and the U.S. hold a commanding lead in AI development, backed by vast financial resources, cutting-edge hardware, and a broad base of human capital to fuel their AI programs.<sup>1</sup> Consequently, the two countries are also fierce competitors in the field of AI.

Within this context of fierce rivalry, the Philippines currently operates within a fragmented landscape of AI-related policy initiatives, lacking a unified national strategy. Technological development remains almost entirely dependent on foreign advancements. The vast majority of AI tools and platforms in use are sourced from international tech firms—many of which are either U.S.-based or linked to Chinese companies whose data practices and security standards raise significant concerns.<sup>2</sup>

Seeing the implications, other countries have made efforts to develop sovereign AI. According to Imperial College London Professor David L. Shrier, “Sovereign AI... is an AI system that is controlled by a government or group of governments, rather than by a single private sector company like OpenAI or Google or Meta.”<sup>3</sup> In fact, Singapore has made commitments and already partnered with NVIDIA to build its own.<sup>4</sup> Vietnam and Thailand have also pursued localized AI initiatives,<sup>5</sup> and it is anticipated that Malaysia will spearhead regional sovereign AI dialogues as the 2025 ASEAN chair.<sup>6</sup>

The Philippines is among the top AI users in the world,<sup>7</sup> but this is as individuals,<sup>8</sup> not in more organized or institutional settings. Should the Philippines fail to evolve a strategy for sovereign AI adoption, we are left

vulnerable as mere end-users of transformative technology. Rather than realizing economic gains because of our talent and high engagement with AI, we might lose higher-value opportunities to other countries. We imagine that AI will continue to drastically transform society, markets, and the way that work is valued and structured. Without a national and sovereign approach to this phenomenon, then we would be allowing the economic and technological “winners” of the AI race, likely the U.S. and China, to dictate what the terms of this new world order would be.

There is already research published that indicates how AI systems are contributing to new global inequalities. For instance, Filipino workers are among those performing essential but underpaid and undervalued jobs, such as data labeling, that are essential to training AI models.<sup>9</sup> The digital divide that currently exists would be widened as access to economic opportunity and education will be limited.<sup>10</sup>

It becomes imperative then that the Philippines develops an agenda focusing on sovereign AI. Such an approach ensures that the Philippines takes control of its future, rather than being dependent and allowing foreign tech elites or hegemonic states to largely influence its economic, technological, and geopolitical prospects.<sup>11</sup>

### **Elevating the strategy**

However, much of the discourse on digital governance and development in the Philippines focuses on data sovereignty—a concept often conflated with the physical localization of data within national borders. Data sovereignty centers on who governs the data and under which laws. Data localization dictates where data must be physically stored and processed as an implementation of sovereignty.<sup>12</sup> Lobbyists and stakeholders in the technology and infrastructure sectors have aggressively debated data localization as a foundational requirement for digital independence.<sup>13</sup> While this approach appears to offer a form of national control, it risks becoming a superficial fix that distracts from the more profound and urgent need for sovereign AI.

Proponents of data sovereignty argue that data localization can improve security, safeguard privacy, ensure compliance with local laws, and enable

economic development by stimulating the construction of data centers.<sup>14</sup> These arguments are not without merit, but risk oversimplifying the AI value chain. Data alone does not confer power. Without the ability to develop, fine-tune, and deploy models on top of that data, a country remains a source of raw material and not an architect of its digital and technological destiny.

Moreover, the risks of focusing solely on data localization are considerable. First, it can lead to complacency, creating the illusion of sovereignty while key AI infrastructure and algorithms remain controlled by foreign entities.<sup>15</sup> Second, it may impose significant cost burdens on businesses and startups, especially when not matched with local computing or AI capacity.<sup>16</sup> Third, localized data is still vulnerable to surveillance, extraction, and manipulation if processed through foreign-owned platforms or APIs. In extreme cases, overreliance on data localization may invite regulatory capture by corporate interests without materially improving national resilience and competitiveness.

Therefore, the Philippines should embrace sovereign AI as a comprehensive, strategic, and future-proof approach for local infrastructure and AI development. By shifting the narrative from where data is stored to how intelligence is generated and owned, we create stronger justifications for investing in national data centers, energy infrastructure, compute capacity, and local AI talent pipelines. Data center utilization should be driven not merely by compliance requirements alone but by a vision of local AI innovation, model training, and responsible deployment.

Sovereign AI thus moves beyond the reactive logic of data localization and toward an affirmative agenda of technological agency, economic inclusion, and geopolitical autonomy.<sup>17</sup> Investing in local talent, research, and infrastructure allows the Philippines to tailor AI solutions to Filipino needs, cultural values, and languages. This mitigates dependency on foreign-developed AI, reducing risks like bias and misalignment with local context.<sup>18</sup> By developing its own AI capabilities, the Philippines can also protect critical sectors, such as, defense, public services, and disaster response, from undue reliance on foreign technology, reducing exposure to external disruptions or security risks.<sup>19</sup> Moreover, sovereign AI can ensure that AI systems reflect Philippine laws, social norms, and language diversity, preventing the

import of alien biases or geopolitical interests embedded in foreign models.<sup>20</sup>

## A dual approach

Within this context, the first step toward Philippine AI Sovereignty will be governance at the national level. At present, the Philippines' approach to AI is a patchwork of roadmaps and agendas, as well as pending bills and other initiatives. While it is encouraging to see evidence of many stakeholders recognizing the importance and pushing for the adoption of an AI strategy, all of this interest merely proves the point that a governing body must be operating at the national level with the resources to plan, coordinate, and execute a sovereign AI plan. The immense scale of AI, along with the various specializations and skills necessary to harness it effectively, is a fundamental roadblock in the execution of AI sovereignty. As such, realizing this framework demands a multidisciplinary approach involving various departments and offices, like the obvious ones in charge of information technology, science, economic development, and trade, but also others that do not appear directly related but are crucial in determining the direction and approach to AI, such as education and human rights. At this point, what we see is that decisive leadership and political will must be exerted toward coordinating this effort, much like a general leading the push on multiple battlefronts.

Another challenge to creating sovereign AI is the lack of the necessary infrastructure to develop it. Additionally, the Philippines has the second most expensive power costs in ASEAN, surpassed only by Singapore.<sup>21</sup> Reliable internet connection also remains a persistent issue. As a result, while there are data centers in the Philippines, these are expensive to maintain and operate.<sup>22</sup>

Furthermore, if we only look at leading AI companies like OpenAI, Anthropic, and Google that are building foundational AI models, then the estimated cost of investment ranges from US\$50 million to US\$100 million.<sup>23</sup> In parallel, NVIDIA sells chips, which it promotes as foundational to AI sovereignty, empowering countries with the necessary computing power. However, developing and deploying such systems would cost hundreds of millions of dollars.<sup>24</sup>

Interestingly, the Chinese AI firm DeepSeek has shown more cost- and energy-efficient ways to build

foundational models. Based on their published information, DeepSeek supposedly cost US\$6 million to develop.<sup>25</sup> This is by no means a small amount, but it is a fraction of the cost of the models developed in the United States.<sup>26</sup>

Apart from infrastructure and engineering, AI education and training at various levels are also needed. The focus has largely been on college-level and youth programs, but we propose that all levels and ages be considered. Educating on AI at all levels should lead to higher AI usage and deployment, which should generate economic value and leverage points for us. The Philippines would have the human capital essential to build sovereign AI and a populace empowered to utilize, help train, and help evolve AI, making it a worthwhile investment. Moreover, we should also look at upskilling programs for people whose jobs will be displaced by AI and facilitate business transformation through the training of business leaders and executives.

Identifying leadership, building infrastructure, as well as facilitating and improving training and education programs are all critical for sovereign AI. Nevertheless, progress will be dependent on political will and the effective development of national policies and frameworks. Ultimately, we will need a policy and regulatory regime that will create and foster the space to build, deploy, and constantly maintain and evolve our sovereign AI.

At the same time, national efforts to realize AI sovereignty can also be strengthened and complemented by strategic international cooperation and partnerships. One of the major challenges to developing AI sovereignty is the country's perceived lack of influence and leverage in terms of AI development.<sup>27</sup> Still, the Philippines is not the only country upon whom this narrative is thrust. That is why it is essential that the Philippines not only imagines itself as having the capacity for AI sovereignty, but also that it joins and advances regional alliances of nations with sovereign AI agendas.

There are two levels at which this can be immediately impactful. First is in terms of building leverage. Whatever body or mechanism is established to bind international partners in advancing their sovereign AI programs, the core idea is that a unified front would carry greater leverage and enable more relatively balanced negotiations with more powerful AI powers,

particularly China and the United States. As is already known, China and the U.S. maintain superior leads in AI development. However, these leads can be overcome or managed with the right moves and investments. Smaller AI states might struggle to achieve competitiveness, secure favorable economic terms, advance their AI agendas, or defend against AI-related threats on their own, but international cooperation opens a pathway to enhance their strategic position and presence in the global AI landscape.

International cooperation also opens the possibility of sharing infrastructure and other resources across the region or with other like-minded partners. Within ASEAN, Singapore has already started a Sovereign AI program, and Malaysia,<sup>28</sup> Vietnam,<sup>29</sup> and Thailand<sup>30</sup> have or are building data centers for AI. These leads and resources could be the starting points for regional collaboration in terms of sharing infrastructure, research, and other resources. By combining resources across states, we can accelerate the progress of AI systems and sovereign AI for each state. The Philippines can and should lead in the formation of this regional body and emphasizing AI as a specific focus of the ASEAN. This would create a stronger negotiating position and a stronger infrastructure and development environment for ASEAN member states.

## Conclusion

At many points in its history, the Philippines has often been on the receiving end of new technologies and the world order that these advancements create. As the world continues to advance and change with AI, there is a possibility that the Philippines will remain a passive recipient of this technology. However, we must stress that this is one front in which we cannot fail to act. At the risk of missing out on political and economic opportunities, the Philippines must begin to imagine itself as a player in the development, implementation, and ongoing evolution of AI as a transformative technology. From this position, we should evolve a national strategy for AI that is anchored on the idea that we must pursue sovereign AI.

Once we begin to advance a unified and national strategy toward sovereign AI, we can build regulatory frameworks to advance AI adoption.<sup>31</sup> Systematic adoption will lead to economic progress, which will

then yield more geopolitical leverage. Unfortunately, the current narratives remain too focused on the AI superpowers. However, nations like the Philippines must be involved in the future development of AI, and we are at a key moment of opportunity to be leaders in the AI revolution.

This article provided a nuanced examination of the Philippines' position in the rapidly evolving landscape of artificial intelligence, highlighting the global dominance of China and the U.S. as AI superpowers. While these two nations lead in AI investment, innovation, and infrastructure, the Philippines remains a significant end-user, highly engaged at the individual level but lacking a coordinated, national strategy for AI development. Currently, Philippine AI policy is fragmented, over-reliant on foreign technology, and focused narrowly on data sovereignty, often conflating it with data localization. While important, this focus alone risks only superficial control and does not address deeper dependencies on foreign AI infrastructure and algorithms. Merely storing data locally will not empower the Philippines to shape its digital future; instead, it risks perpetuating the nation's status as a raw data provider and an undervalued contributor to the AI value chain.

To counteract this, we strongly advocate for a shift toward Sovereign AI, the creation of local capacity to build, deploy, and control AI systems tailored to the country's specific needs, values, and contexts. The benefits are compelling: reducing technological dependency, protecting critical sectors, aligning AI with Filipino culture and law, and providing higher-value economic opportunities rather than relegating workers to low-wage, low-skill roles in the global AI supply chain.

Significant barriers exist. Nevertheless, if the Philippines wishes to avoid technological dependency and economic marginalization in the AI era, it must move urgently beyond piecemeal policies and superficial data localization. A bold national agenda for sovereign AI is required, one that prioritizes local talent, research, and infrastructure; fosters inclusive, responsible innovation; and seeks out strategic international alliances. Only through such a comprehensive and forward-thinking strategy can the Philippines secure its technological agency, economic competitiveness, and geopolitical autonomy in the coming age of artificial intelligence.

## References

- 1 R. Preston McAfee, Review: Review of *AI Superpowers China, Silicon Valley and the New World Order*, by Kai-Fu Lee," *Business Economics* 54, no. 3 (2019): 185-190, <https://www.jstor.org/stable/48755677>.
- 2 Jose Miguelito Enriquez, "Philippine AI Governance: Time to Shift Gears," *Fulcrum*, April 16, 2025, <https://fulcrum.sg/philippine-ai-governance-time-to-shift-gears/>.
- 3 David L. Shrier, "Building the Future Federation - The Rising Calls for Sovereign AI," *CIRSD Horizons Summer 2024*, no. 27 (2024), <https://www.cirsd.org/en/horizons/horizons-summer-2024--issue-no-27/building-the-future-federation---the-rising-calls-for-sovereign-ai>.
- 4 Joao Gomes, "Singtel, NVIDIA to Bring Sovereign AI to Southeast Asia," *NVIDIA*, January 30, 2024, <https://blogs.nvidia.com/blog/singtel-sovereign-ai/>.
- 5 Bama Andika Putra, "Governing AI in Southeast Asia: ASEAN's way forward," *Frontiers in Artificial Intelligence* 7, (2024): 1-6, <https://doi.org/10.3389/frai.2024.1411838>.
- 6 "Strengthening Regional Cooperation," About Us, Sovereign AI Project, accessed August 2025, <https://sovereign-ai.org/about/>.
- 7 Pia Piquero, "Philippines among top 5 countries for ChatGPT usage, report shows," *Cebu Daily News*, October 23, 2024, <https://cebudailynews.inquirer.net/601869/philippines-among-top-5-countries-for-chatgpt-usage-report-shows>.
- 8 Microsoft Philippines Communications Team, "Microsoft and LinkedIn Release 2024 Work Trend Index on AI Use at Work in the Philippines," *Microsoft*, May 23, 2024, <https://news.microsoft.com/en-ph/2024/05/23/microsoft-and-linkedin-release-2024-work-trend-index-on-ai-use-at-work-in-the-philippines/>.
- 9 "AI underpinned by an invisible and exploited workforce," *Svenssonstiftelsen*, October 15, 2024, <https://www.svenssonstiftelsen.com/post/ai-underpinned-by-an-invisible-and-exploited-workforce>.
- 10 UP Media and Public Relations Office, "Nurturing the Next Generation of Leaders who will navigate and shape AI's future," *University of the Philippines*, May 13, 2025, <https://up.edu.ph/nurturing-the-next-generation-of-leaders-who-will-navigate-and-shape-ais-future/>.
- 11 "The Rise of Sovereign AI: National Strategies, Global Implications," *Alphanome.AI*, April 27, 2025, updated May 14, 2025, <https://www.alphanome.ai/post/the-rise-of-sovereign-ai-national-strategies-global-implications>.
- 12 Mesh Flinders and Ian Smalley, "What is data sovereignty?," *IBM*, June 3, 2024, <https://www.ibm.com/think/topics/data-sovereignty>.
- 13 Rainier Allan Ronda, "Business, IT groups nix plans for data localization," *Philstar*, August 27, 2023, <https://www.philstar.com/business/2023/08/27/2291580/business-it-groups-nix-plans-data-localization>.
- 14 Danielle Barbour, "Data Sovereignty: a Best Practice or Regulatory Requirement?," *Kiteworks*, July 22, 2024, <https://www.kiteworks.com/regulatory-compliance/data-sovereignty-best-practice-regulatory-requirement/>.
- 15 "The 'Real Life Harms' of Data Localization Policies," *Centre for Information Policy Leadership Discussion Paper* 1, (2023): 1-14, [https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-tls\\_discussion\\_paper\\_paper\\_i\\_-\\_the\\_real\\_life\\_harms\\_of\\_data\\_localization\\_policies.pdf](https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-tls_discussion_paper_paper_i_-_the_real_life_harms_of_data_localization_policies.pdf).
- 16 "Data localization requirement: The Impact of Data Localization on Marketing Campaigns for Startups," *FasterCapital*, April 11, 2025, <https://www.fastercapital.com/content/Data-localization-requirement--The-Impact-of-Data-Localization-on-Marketing-Campaigns-for-Startups.html>.
- 17 Dan Rowinski, "World Leaders Put Spotlight on Sovereign AI at NVIDIA GTC," *NVIDIA*, March 4, 2025, <https://blogs.nvidia.com/blog/sovereign-ai-gtc-2025/>.
- 18 Muath Alduhishy, "Sovereign AI: What it is, and 6 strategic pillars for achieving it," *World Economic Forum*, April 25, 2024, <https://www.weforum.org/stories/2024/04/sovereign-ai-what-is-ways-states-building/>.
- 19 Serene Nah, "How sovereign AI will reshape APAC's digital future," *ComputerWeekly.com*, May 13, 2025, <https://www.computerweekly.com/opinion/How-sovereign-AI-will-reshape-APACs-digital-future>.
- 20 "Sovereign AI Explained: Power, Privacy & Potential," *daffodil*, July 17, 2025, <https://insights.daffodilsw.com/blog/sovereign-ai-explained-power-privacy-potential>.
- 21 "PH power prices among highest in ASEAN, next to Singapore," *Power Philippines*, September 19, 2023, <https://powerphilippines.com/ph-power-prices-among-highest-in-asean-next-to-singapore/>.
- 22 "Saavedra: High power costs and connectivity issues keep Philippines behind in data centers," *Bilyonaryo*, January 23, 2025, <https://bilyonaryo.com/2025/01/23/saavedra-high-power-costs-and-connectivity-issues-keep-philippines-behind-in-data-centers/technology/>.
- 23 Ben Cotteir, Nestor Maslej, Robi Rahman, Loredana Fattorini, and David Owen, "The rising costs of training frontier AI models," *arXiv*, May 31, 2024, <https://arxiv.org/html/2405.21015v1>.
- 24 Albert Fox, "The AI Infrastructure Gold Rush: NVIDIA's \$500 Billion Gamble and Its Investment Implications," *AlInvest*, April 15, 2025, <https://www.ainvest.com/news/ai-infrastructure-gold-rush-nvidia-500-billion-gamble-investment-implications-2504/>.
- 25 Ben Wodecki, "Behind the DeepSeek hype: Costs, safety risks & censorship explained," *Capacity*, January 29, 2025, <https://www.capacitymedia.com/article/2ecgy5isrr4777k2yws1s/long-reads/behind-the-deepseek>.
- 26 "How energy-efficient is DeepSeek, China's AI disruptor?," *Rinnovabili*, January 29, 2025, <https://www.rinnovabili.net/business/markets/deepseeks-energy-consumption-ais-75-power-cut/>.
- 27 Swati Srivastava and Justin Bullock, "AI, Global Governance, and Digital Sovereignty," *arXiv*, October 23, 2024, <https://arxiv.org/html/2410.17481v1>.
- 28 Reuters, "Microsoft Malaysia to launch three data centres in Q2 2025," *Reuters*, March 20, 2025, <https://www.reuters.com/technology/microsoft-malaysia-launch-three-data-centres-q2-2025-2025-03-20/>.
- 29 Samaya Dharmaraj, "AI Demand Fuels Data Centre Growth in Vietnam," *Taipei Economic and Cultural Office in Vietnam*, February 28, 2025, [https://www.roc-taiwan.org/vn\\_en/post/5983.html](https://www.roc-taiwan.org/vn_en/post/5983.html).
- 30 "New data centre rises in Bangkok to meet growing AI needs," *The Nation*, November 26, 2024, <https://www.nationthailand.com/business/tech/40043614>.
- 31 Dominic Vincent Ligot, "AI Governance: A Framework for Responsible AI Development," *SSRN*, May 5, 2024, <https://dx.doi.org/10.2139/ssrn.4817726>.

# Author Profiles

---

## Joshua Uel Abad

Program Manager for Inclusive Economic Development  
KAS Philippines

Joshua Uel Abad is currently the Program Manager for Inclusive Economic Development at KAS Philippines. With over 8 years of research and project management experience in various fields such as Workforce Development, Education, and Economic Regulation, he has successfully managed and implemented a variety of programs that drive sustainable growth and economic empowerment.

He started his career with the Asian Institute of Management, where he managed policy research projects and facilitated dialogue between researchers, government stakeholders, and civil society organizations. Before joining KAS, he co-managed a million-dollar technical assistance program funded by the Australian Government to address the workforce development and educational gaps brought about by the 4th Industrial Revolution in various sectors in the Philippines.

He holds a Bachelor's degree in Political Economy from the University of Asia and the Pacific. His research interests include economic regulation, workforce development, education policy, governance innovation, public sector reform, and evidence-based policymaking.



## Nathan Samiley

Deputy Executive Director  
AAP

Nathan Samiley is currently the Deputy Executive Director of the Analytics and Artificial Intelligence Association of the Philippines (AAP). In this role, he focuses on strengthening AAP's position as an ecosystem builder, fostering greater collaboration among stakeholders to achieve the AAP's goal of sustaining the analytics and AI community in the Philippines.

He has 13 years of combined work experience in the non-profit, government, and private sectors. His previous stints included being the Partner Programs and Sustainability Lead for Data Analytics Ventures, Inc. (DAVI), Membership Lead of the German-Philippine Chamber of Commerce and Industry (GPCCI), the Philippine Software Industry Association (PSIA), and the Animation Council of the Philippines (ACPI). As part of his government experience, he was with the Design Center of the Philippines, an attached agency of the Department of Trade and Industry (DTI), under its Planning, and eventually, the Design Promotion Office. At the DTI, he led various events and promotional activities, including the Design Week Philippines and the ASEAN Creative Cities Forum and Exhibition, one of the key events of the DTI for the ASEAN's 50th Anniversary.



## **Eylla Laire M. Gutierrez, Ph.D.**

OIC Executive Director

AIM-Dr. Andrew L. Tan Center for Tourism

Dr. Eylla Laire M. Gutierrez is a development consultant and the OIC Executive Director of the Asian Institute of Management's Dr. Andrew L. Tan Center for Tourism. She also serves as a fellow of the Ryoichi Sasakawa Young Leaders Fellowship (SYLFF) Japan. She has completed her Doctor of Philosophy (Ph.D.) degree in Asia Pacific Studies at the Ritsumeikan Asia Pacific University, Japan, and was also conferred with the Outstanding Dissertation Award. Her research interests include sustainability, community development, gender studies, and tourism management.



## **Ray Justin A. Villanueva**

Program and Research Associate

AIM-Dr. Andrew L. Tan Center for Tourism

Ray Justin A. Villanueva is the Program and Research Associate at the Dr. Andrew L. Tan Center for Tourism, Asian Institute of Management. He graduated summa cum laude from the University of the Philippines Diliman, earning a Bachelor's degree in Economics with an informal minor in Tourism Management and the distinction of College Valedictorian.

His research interests span tourism marketing, value chain analysis, local tourism campaigns, community-based tourism policies, and digitalization in the tourism industry. His work at the Center focuses on bridging academic research and practical application by organizing tourism events and dialogues, as well as facilitating workshops on tourism development planning in local communities.



## Marie Antoinette P. de Jesus

Senior Program Manager for Governance and Foreign Relations  
KAS Philippines

Marie Antoinette P. de Jesus is a researcher and writer specializing in governance, foreign policy, and security issues. As Senior Program Manager for Governance and Foreign Relations at the Konrad-Adenauer-Stiftung (KAS) Philippines, she develops research programs, edits publications, and writes extensively on governance, political economy, and maritime security, among other topics. Her work connects policy research and public discourse, with essays and analyses published by Rappler, 9DashLine, and *The KAS Solutions Review* series. With experience spanning the World Bank Office in Manila, the Asian Development Bank Knowledge Hub at the Asian Institute of Management, and the International Rice Research Institute, she gives a research-driven lens to multi-layered policy questions.



## Jaemark Tordecilla

Independent Journalist and Media Consultant  
Philippines

Jaemark Tordecilla is a journalist, media advisor, and technologist from the Philippines, specializing in newsroom innovation and artificial intelligence. He is currently working with newsrooms and media development networks across Asia on innovation projects, helping news organizations adapt to emerging technologies.

He was a Nieman Fellow at Harvard University in 2023-2024, focusing on the implications of generative AI in the media industry, including its ethical considerations and potential applications. Prior to his fellowship, he spent nearly a decade as editor-in-chief and head of digital media at GMA News Online. In 2021, he was honored with The Outstanding Young Men (TOYM) Award, one of the Philippines' top honors for young civic leaders, in recognition of his contributions to digital journalism.

His work on building AI tools for newsrooms in developing countries has been featured by the Reuters Institute and the Computational Journalism Lab at Northwestern University. His research and insights on AI in journalism have also been cited in key industry reports, including the Reuters Institute's Digital News Report 2024, WAN-IFRA's Innovation in News Media World Report 2024-25, and the Gen AI x Comms Industry Impact Report 2024. He has since conducted dozens of seminars and workshops on generative AI and digital newsroom strategy and transformation for journalists and media managers from across Southeast Asia, the United States, South Asia, and Latin America.



## Carljoe Javier

Executive Director and Chief AI Ethicist  
Data and AI Ethics PH

Carljoe Javier is the Executive Director and Chief AI Ethicist of Data and AI Ethics PH, a company dedicated to advancing responsible and ethical use of data and artificial intelligence. He previously served as CEO of *PumaPodcast*, an award-winning Philippine podcast production company. He was the Managing Editor of Anino Comics and author of books of fiction, nonfiction, and comics. His latest book, *Pop Culture & AI*, dives into the hopes, fears, and ethical trade-offs of our tech-driven future.



## Dominic Vincent Ligot

Founder  
CirroLytix and Data and AI Ethics PH

Dominic Vincent "Doc" Ligot is one of the leading voices in artificial intelligence in the Philippines, known for his pioneering work in AI ethics, responsible AI adoption, and social impact through data. He is the founder of CirroLytix, an award-winning social impact AI training and consulting company, and AI and Data Ethics PH, an advocacy community tackling the societal implications of data and AI. He also serves as Chair of AI Ethics and Safety at the Philippine AI Business Association.

Doc has been extensively cited in local and global media outlets, including *The Economist*, *South China Morning Post*, *Washington Post*, and *Agence France-Presse*, for his expertise in AI and data governance. His award-winning work has been recognized and published by prestigious organizations such as NASA, Data.org, Digital Public Goods Alliance, the Group on Earth Observations (GEO), the United Nations Development Programme (UNDP), the World Health Organization (WHO), and UNICEF.

He is a member of the Expert Advisory Panel for the International Scientific Report on Advanced AI Safety, convened by the UK Department of Science, Innovation, and Technology. He also co-authored the Master in Applied Business Analytics degree at the University of Asia and the Pacific. He was also instrumental in leading Project SPARTA, a nationwide data science education initiative in the Philippines.



# The Editors

---

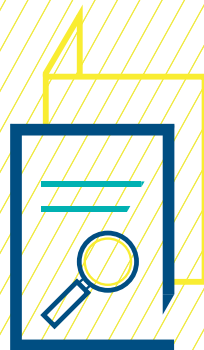
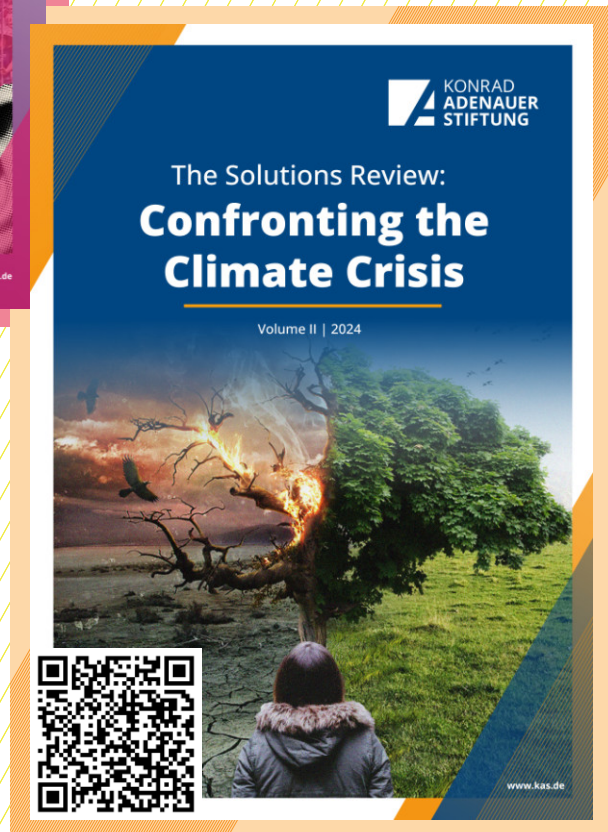
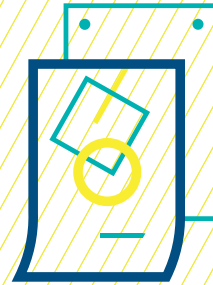
**Sophiya Navarro** is the Program Manager for Security and Innovation at KAS Philippines and Managing Editor of *The Solutions Review* publication series. Her professional and research interests revolve around exploring the intersections between good governance, security (especially maritime security), and innovation. Before joining KAS, she was a senior research and international cooperation specialist at the Philippine Space Agency (PhilSA) from 2020 to 2022 and Program Coordinator and Research Assistant at the Ateneo de Manila University – School of Government from 2018 to 2020. She holds a Bachelor’s degree in Diplomacy and International Relations with a specialization in East and Southeast Asian Studies and a Minor in Literature (English) from the Ateneo de Manila University, where she graduated with honors in 2018 under a full scholarship.

**Frank Cally Tabuco** is a freelance Machine Learning Engineer and AI researcher with expertise spanning computer vision, signal processing, NLP, and applied deep learning. He has over four years of experience in developing and deploying AI solutions for industries and research, including applications in sleep staging, thunderstorm prediction, heart failure detection, and behavioral malware analysis. His work focuses on building practical, high-impact AI systems that address real-world challenges. He holds a Master of Science degree in Computer Science from the University of the Philippines and a Bachelor of Science degree in Management Engineering from the Ateneo de Manila University.

**Daniela Braun** is the Country Director of KAS Philippines and interim Country Director of KAS Cambodia. Previously, she worked as a Policy Advisor for Foreign and Security Affairs at the KAS Headquarters in Berlin and a Program Officer at the German Council on Foreign Relations (DGAP). She holds a Master’s degree in political science from the Free University Berlin (2014) and a Bachelor’s degree from the Rheinische-Friedrich-Wilhelms-University in Bonn, Germany (2011). She was a Fox Internal Fellow at Yale University from 2016-2017.

# Solutions begin with us.

Catch up on past issues, stay engaged, and help shape the future of democracy.



# The Konrad-Adenauer-Stiftung (KAS)

KAS is a German political foundation and think tank, closely associated with the Christian Democratic Union of Germany (CDU). As co-founder of the CDU and the first Chancellor of the Federal Republic of Germany, Konrad Adenauer (1876-1967) united Christian-social, conservative and liberal traditions. His name is synonymous with the democratic reconstruction of Germany, the firm alignment of foreign policy with the trans-Atlantic community of values, the vision of a unified Europe, and an orientation towards the social market economy. His intellectual heritage continues to serve both as our aim and as our obligation today.

In our European and international cooperation efforts, we work for people to live self-determined lives in freedom and dignity. We make a contribution underpinned by values to help Germany meet its growing responsibilities throughout the world. We encourage people to contribute to shaping the future. With more than 70 offices abroad and projects in over 120 countries, we uniquely contribute to promoting democracy, the rule of law, and a social market economy. To foster peace and freedom, we encourage continuous dialogue at the national and international levels and the exchange between cultures and religions.

KAS has been active in the Philippines since 1964. It is the oldest KAS office in Asia. Our pillars of work in the Philippines are promoting democracy and good governance, inclusive economic development, foreign relations, security, innovation, and youth empowerment.

Our vision is to provide a safe space for everyone to discuss politics and give a voice to people who need a platform. We cooperate with government institutions, political parties, civil society organizations, experts, and key stakeholders, building strong partnerships along the way to uphold democracy and provide a platform for issues close to our hearts. Together with our partners, we provide research-based policy inputs and create momentum for relevant policy pushes.

Follow KAS Philippines online through the following channels:



[kas.de/philippines](https://kas.de/philippines)



[facebook.com/kasphilippines/](https://facebook.com/kasphilippines/)



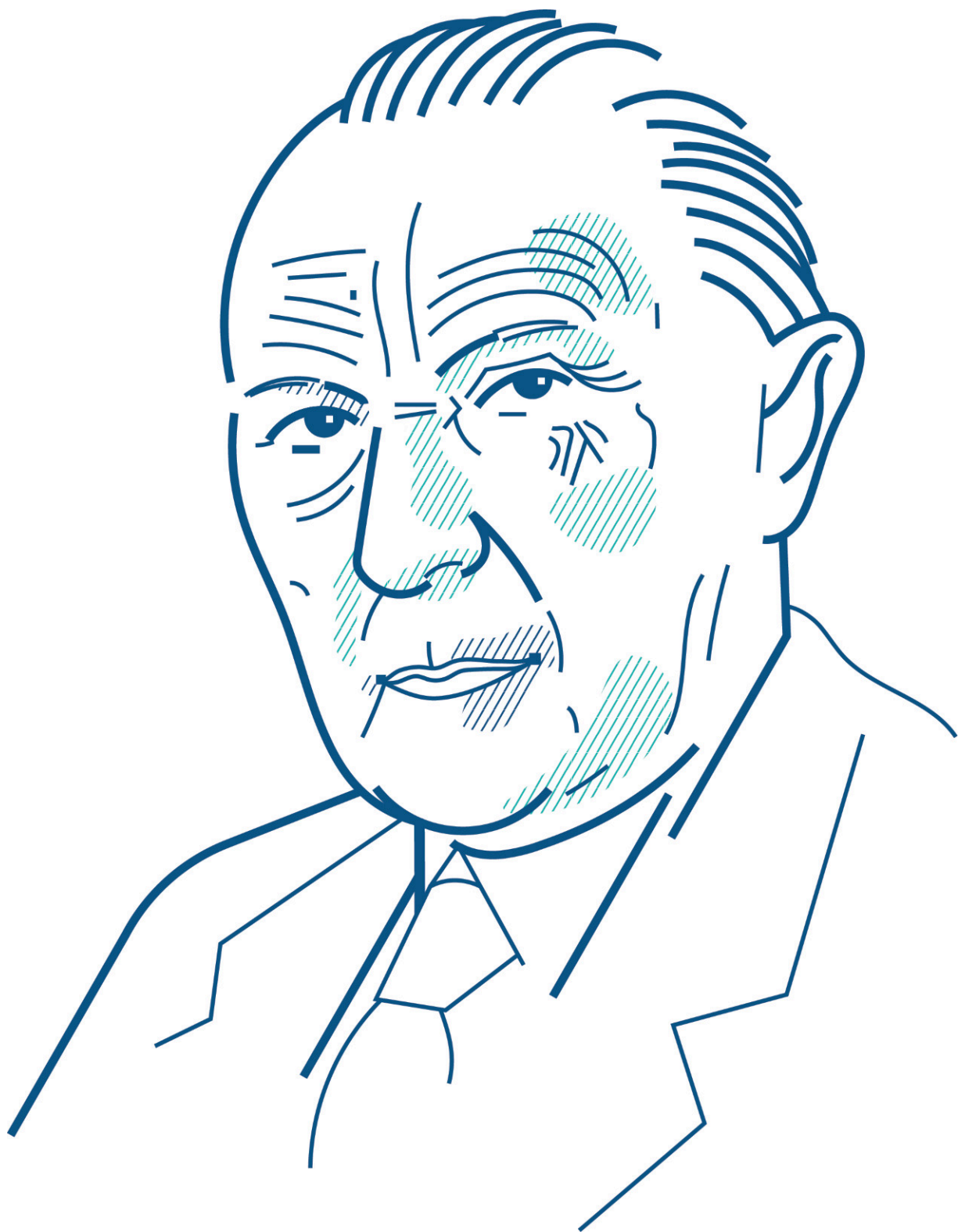
[instagram.com/kasphilippines/](https://instagram.com/kasphilippines/)



[x.com/kasphilippines](https://x.com/kasphilippines)



[linkedin.com/company/kasphilippines/](https://linkedin.com/company/kasphilippines/)





**ISSN**  
3116-2738 (online)  
3116-272X (print)