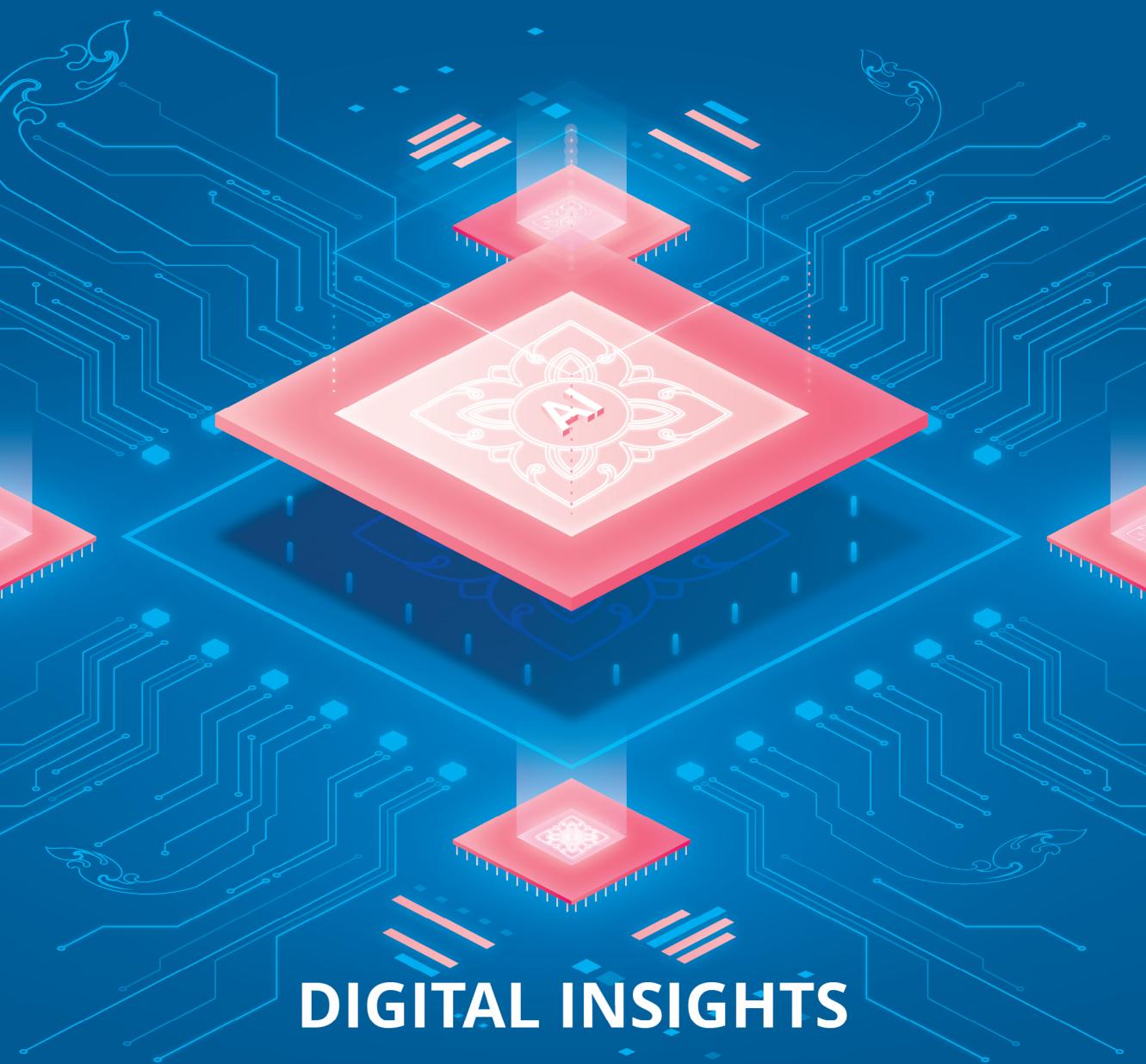


# The Future of AI for Development



**DIGITAL INSIGHTS**

**2026, Konrad-Adenauer-Stiftung Cambodia**

**Edited by**

Pisal CHANTY, Jonas HERZING, Mouy Ing SEAN

**Production Team**

Mouy Ing SEAN, Sokun Kosoma KOEUT, Sokleap LIM

**Cover and Layout Design**

Brosnuth SOEUN

---

**Konrad-Adenauer-Stiftung Cambodia**

House No. 4, Street 462, Khan Charmkar Mon, P.O. Box 944,  
Phnom Penh, Kingdom of Cambodia

Telephone: +855 23 966 176

Email: [Office.Phnompenh@kas.de](mailto:Office.Phnompenh@kas.de)

Website: [www.kas.de/cambodia](http://www.kas.de/cambodia)

Facebook: [www.facebook.com/kaskambodscha](http://www.facebook.com/kaskambodscha) Printed in Cambodia

*Printed in Cambodia*

*ISBN: 978-9924-571-36-0*



Cambodia Academy of Digital Technology (CADT) is a national research and education institution for Digital Technology and Innovation. Its mission is to provide high-quality higher education and professional training in digital technology to students, professionals, and civil servants who want to excel in the digital economy and society. CADT is the first institution of its kind in Cambodia to focus on research and development in digital technologies that advances knowledge and creates value from digital adoption and transformation. CADT fosters a culture of innovation and entrepreneurship among digital talents and innovators to support academic and socio-economic development. CADT was established in 2014 as the National Institute of Post, Telecom, and ICT (NIPICT), a public institution under the Ministry of Post and Telecommunications. In 2021, it transformed into CADT with an expanded mission to support the country's digital transformation agenda and the development of digital economy and society.



Freedom, justice and solidarity are the basic principles underlying the work of the Konrad-Adenauer-Stiftung (KAS). The KAS is a political foundation, 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 transatlantic 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 as well as our obligation today.

In our European and international cooperation efforts we work for people to be able to live self-determined lives in freedom and dignity. We make a contribution underpinned by values to helping Germany meet its growing responsibilities throughout the world.

KAS has been working in Cambodia since 1994, striving to support the Cambodian people in fostering dialogue, building networks and enhancing scientific projects. Thereby, the foundation works towards creating an environment conducive to economic and social development. All programs are conceived and implemented in close cooperation with the Cambodian partners on central and sub-national levels.

*This page is intentionally left blank.*

# **DIGITAL INSIGHTS**

**01/2026**

# FOREWORD

---

## Daniela Braun

Country Director

Konrad-Adenauer-Stiftung (KAS) Cambodia

Cambodia stands at a pivotal moment in its digital transformation journey. Although still ranked low on a global and regional scale, the country has made progress in expanding connectivity, strengthening its digital infrastructure, and fostering an increasingly dynamic technology ecosystem. The next phase of this journey will be significantly defined by the strategic adoption of emerging technologies—most notably AI. It is within this context that *Digital Insights*, a two-part publication by CADT and KAS Cambodia, offers a timely and policy-relevant contribution to the national conversation on digital development.

This publication is intentionally designed to depict AI not as an abstract technology, but as a practical and enabling tool with direct relevance to Cambodia's socio-economic development. Rather than taking a generic approach, *Digital Insights* focuses on concrete applications of AI across sectors that are central to the country's long-term growth. By doing so, KAS Cambodia wants to contribute to bridging the gap between technological possibility and policy action, identifying where AI can meaningfully support key national priorities—from food security and agricultural modernization to SME competitiveness, and effective public administration.

One of the core strengths of *Digital Insights* is the careful calibration of its scope. We recognize that effective policymaking requires more than theoretical assessments; it requires applied analysis grounded in local realities. The publication therefore examines specific case studies that are highly relevant for the Cambodian context, while also drawing cross-cutting insights that are essential for designing cohesive national strategies. This balance ensures that this publication is both comprehensive and practical—broad enough to inform multi-sectoral planning, yet focused enough to guide actionable decision-making.

The agricultural sector is a prime example. As a backbone of Cambodia's economy, agriculture faces structural challenges and *Digital Insights* highlights how AI-driven tools can help farmers optimize resources, manage climate risks, and improve market access. For the country's large SME sector, AI offers opportunities to boost efficiency, enhance supply chain management, and support innovation-driven growth. Meanwhile, in governance, the careful use of AI can help modernize public service delivery, streamline administrative processes, and strengthen transparency—provided that robust safeguards and ethical frameworks are in place.

Ultimately, *Digital Insights* serves as a call to action for policymakers, industry leaders, researchers, and development partners. It underscores that AI adoption is not simply a technological choice; it is a strategic imperative for maintaining competitiveness, unlocking new economic opportunities, and ensuring sustainable development.

I would like to express my sincere appreciation to all contributors whose expertise and insights greatly enriched this publication. My thanks also go to our partner, the Cambodia Academy of

Digital Technology (CADT), for their valuable collaboration throughout the process. A special word of gratitude goes to my colleague Mouying Sean, whose dedication and support were instrumental in bringing *Digital Insights* to completion.

I strongly hope that the ideas and insights presented in the following pages will help inform the next steps of Cambodia's digital transformation and contribute meaningfully to sustainable economic growth and improved governance. KAS Cambodia is proud to support this work and remains committed to fostering dialogue and research that drive the country's digital future forward.



### **Daniela Braun**

*Country Director  
KAS Cambodia*

*Daniela Braun is the KAS Country Director Philippines since May 2023 and acting Country Director Cambodia since March 2025. In her capacity, she oversees the programs of KAS in the Philippines and Cambodia on democracy & governance, inclusive & sustainable economic development, security & foreign relations, youth empowerment, media & society as well as digitalization. Previously, Daniela worked as Policy Advisor for International Politics and Security Affairs at the Division Analysis and Consulting of Konrad-Adenauer-Stiftung in Berlin. Before KAS, Daniela held different positions at the German Council on Foreign Relations, the German Bundestag and the Berlin Office of the German Marshall Fund.*

*She holds a B.A. in Political Science from the University of Bonn and M.A. from the Freie University Berlin. She also studied in Warsaw, Poland, at the American University in Washington D.C. and as a Fox International Scholar at Yale University in New Haven.*

# FOREWORD

---

**Dr. Sopheap SENG**

*President*

*Cambodia Academy of Digital Technology (CADT)*

Cambodia stands at a pivotal milestone in its digital transformation journey. With rapid improvements in digital connectivity and infrastructure, the country is laying the foundation for a vibrant digital economy and society. Emerging technologies such as Artificial Intelligence (AI) offer unprecedented opportunities to accelerate this transformation—particularly in key development sectors like education, public health, SME, and agriculture.

Such advanced technologies as AI that have emerged to be one of the powerful and useful technologies in the digital era need thorough understanding with rigorous study and research. As a public education and research institution, CADT is committed to building Cambodia's digital talent and fostering innovation. We focus on practical research which can be applied to enhance existing traditional ways of doing things to create an impactful and sustainable development.

As a public education and research institution, we are dedicated to being a catalyst in driving the development of digital transformation and innovation with a mission and commitment to build Cambodia's resilient digital talent and drive digital research and innovation. Our effort has been highlighted by a recent national achievement in which Cambodia has become the 4th country in Southeast Asia to complete UNESCO's AI Readiness Assessment in collaboration with MPTC. This national report charts a clear path toward ethical, inclusive, and human-centered digital transformation—and contributes to the drafting of the national AI strategy.

Realizing this vision requires collective action. CADT has partnered with various organizations – universities, research centers, and international development organizations, both public and private to ensure the maximization of impact making of digital transformation to the mentioned important sectors. Our partnership with KAS Cambodia is vital because it bridges the gap between theory and practice. By integrating insights from the international development partner with our own evidence-based research, these publications provide readers with a holistic view of Cambodia's current landscape and the academic depth needed to understand its future.

Building on our successful 2023 publication, *The Future of Education*, where digitalization in the education sector was introduced and discussed. CADT and KAS Cambodia are proud to continue this flagship series. This year's edition expands our horizon, exploring how AI can revolutionize key development sectors such as education, public health, and agriculture.

This publication marks a significant milestone of Cambodia's collective and real interest in advanced digital technology like AI regarding the development of the country. We believe this year's edition will be an indispensable resource for educators, researchers, policymakers and innovators to move together toward a digitally advanced and resilient Cambodia through the responsible adoption of AI.



**Dr. Sopheap SENG**

*President*

*Cambodia Academy of Digital Technology (CADT)*

*Dr. Sopheap SENG is the President of the Cambodia Academy of Digital Technology (CADT), a public research and education institution specialized in digital technology and innovation. Dr. Sopheap is also a lecturer in computer science, he teaches programming, database and data science. At CADT he is currently directing the efforts to build a digital innovation center to contribute to the development of the digital startup ecosystem in Cambodia. CADT is pursuing deeper public- private cooperation and is trying to create supportive policies for technology entrepreneurship and innovation. CADT is a key player in digital skill development in Cambodia, nurturing digital talent and innovators and is playing an integral role as the country seeks to formulate ground rules for its budding startup and technology sector.*

## EDITORIAL NOTES

---

Dear readers,

Artificial intelligence in Cambodia is no longer a distant or speculative concept. It now shapes social media feeds, classroom assignments, administrative workflows in ministries, and even how farmers access information on prices and weather. Since the rapid diffusion of generative AI from November 2022 onwards, particularly tools such as ChatGPT, AI has moved from conference presentations into the routines of daily life by building on debates that have, in fact, been present on the government's agenda for many years prior.

Yet if you only followed the global debate, you could easily believe that serious AI strategies belong exclusively to rich countries, while least developed countries (LDCs) are either passive recipients or experimental playgrounds. Cambodia has chosen a different path. Instead of rushing to publish a glossy "AI vision" full of slogans, the Royal Government first subjected itself to a diagnostic: UNESCO's AI Readiness Assessment Methodology (RAM), a dozen indicators toolkit that forces uncomfortable questions about laws, data, skills, infrastructure, and institutions. The assessment concluded in early July 2025, drawing on inputs from 26 ministries and multiple consultation workshops.

The result is a draft National AI Strategy that is unusually grounded for an LDC context. It is explicit about gaps in data, compute, and talent; it admits that Cambodia is at an early stage of AI adoption; and it still argues, unapologetically, that AI must serve economic transformation, not the other way around. The vision is to use AI to lift productivity, expand value creation, and support a more inclusive and sustainable development model – while avoiding heavy-handed regulation that would suffocate a nascent ecosystem before it can grow. One of the six strategic priorities emerging from the RAM process is collaboration, research and development, and innovation. That sounds abstract, until you start asking who is actually talking to whom. Are regulators talking to engineers? Are universities listening to farmers and SMEs? Are global frameworks translated into Khmer and into practice? This book, *The Future of AI for Development*, is our attempt to make that priority visible, concrete, and, we hope, a little bit irresistible to read.

The volume is the product of a partnership between Konrad-Adenauer-Stiftung (KAS) Cambodia and the Cambodia Academy of Digital Technology (CADT). It brings together authors from government, academia, civil society, and the private sector who are directly involved in building Cambodia's AI future. We structured the book deliberately to challenge the usual pattern of "importing" ideas from abroad and "exporting" problems at home. The first part looks outward: it examines international and regional trends in AI governance, from ASEAN's soft-law guidelines to more legalistic models in the EU, Japan, and New Zealand, and asks what actually travels well into the Cambodian context. The second part turns inward, unpacking the readiness assessment and the political, technical, and institutional choices behind Cambodia's National AI Strategy. The third part moves to sectoral applications, such as education, public data, workforce development – before the final part dives into case studies that are already operating on Cambodian soil.

Our editorial stance is simple and intentionally provocative: Cambodia cannot afford to be either blindly optimistic or reflexively fearful about AI. Treating AI as magic will hand power to the few who control the technology; treating it only as a threat will guarantee that others design the systems

that shape our economy, our media, and even our language. The more serious alternative is to treat AI as infrastructure – something that must be planned, regulated, funded, and contested in the open. That means insisting that Khmer language and culture are first-class citizens in training data, that public institutions build real in-house expertise, and that citizens have both the literacy and the channels to question AI systems that affect their lives.

We are deeply grateful to the authors who contributed their time, expertise, and, in many cases, their hard-won field experience; to our colleagues at CADT and KAS Cambodia who supported the project from concept note to printed pages; and to the institutions that opened their data, labs, and classrooms so that AI could be discussed in the light, not in the shadows.

If this editorial note has already kept you longer than you expected, our work here is done. We invite you now to turn the page, not to admire AI from a distance, but to argue with it, test it, and decide what role it should play in Cambodia's development story.

Phnom Penh, December 2025

With warm regards,  
Chanty Pisal  
Jonas Herzing  
Mouy Ing Sean



### Chanty Pisal

*Pisal is the Director of the Digital and Innovation Policy Research Center at the Cambodia Academy of Digital Technology (CADT). He holds a Master of Public Policy from the Australian National University and a Master's degree in International Law from the University of Paris 8. Pisal has contributed to major national digital initiatives and policies, including the Digital Skills Development Roadmap and the Digital Government Policy. He is also a key contributor to Cambodia's AI initiatives, notably the UNESCO Ethics of AI Readiness Assessment.*

### Jonas Herzing

*With over a decade of experience in the field of digital transformation and positions in the private and public sector, Jonas is currently working as Senior Advisor at the GIZ Digital Transformation Center supporting the digital transformation of the economy and government in Cambodia. He holds a double master's degree in international politics with a focus on global digital governance (Fudan University, China and Sciences Po Paris, France) and a bachelor's degree in IT Management (Berlin School of Economics and Law).*



### Mouy Ing Sean

*Mouy Ing is the Program Manager for Global Affairs and Governance at KAS Cambodia. She holds a Bachelor's degree in International Relations from Paragon International University and specializes in research across the social sciences, working to bridge knowledge with real-world impact.*



# CONTENT

## *International and Regional Trends*

**12** Towards Ethical AI Governance on Digital Communication Platforms in Cambodia: Comparative Insights from the ASEAN, EU, Japan, and New Zealand  
*Leavsovath In and Phyrobish Ream*

**26** Operationalizing ASEAN's AI Governance Guidelines: A Policy Implementation Study in Cambodia  
*Catherine Setiawan*

## *Strategy for the Development of AI*

**38** Readiness Assessment and the Development of National AI Strategy in Cambodia  
*Chanty Pisal and Bunhong Taing*

## *The Application of AI in Different Sectoral Development*

**52** AI in Higher Education: Attitudes and Acceptance Among University Students in Cambodia  
*Borey Be*

**78** AI Meets Public Data: A Prompt-Based Agent for Intuitive Analysis of Government Datasets  
*Dr. Rina Buoy, Dr. Nguonly Taing, and H.E. Dr. Marry Kong*

**88** Training the Trainers: Scaling AI Annotation Through Inclusive Workforce Development  
*Songheng Tieng, Sokheng Pok, and Chantola Khet*



## **Case Studies**

**100** Machine Translation System for Khmer Text to Khmer Braille  
*Kimhuoy Yann, Nab Mat, and Rottana Ly*

**110** Reducing Antimicrobial Resistance (AMR) via Improving Microbiology Analysis Capacity and Optimizing the Use of Antibiotics through Machine Learning with Artificial Intelligence (AI)  
*Sophal Thear, Dr. Stanislas Rebaudet, Dr. Kennarey Seang, and Dr. Sotharith Bory*

**122** HealthyRice: Enhancing Farmer-Expert Collaboration via Mobile Technology  
*Hongly Va and Soklong Him*

**134** Policy-Driven Approaches in AI-Powered Verification for Combating Mis- and disinformation: A Cambodia-Focused Study  
*Prohim Tam and Riccardo Corrado*





# INTERNATIONAL AND REGIONAL TRENDS

**12**

Towards Ethical AI Governance on Digital Communication Platforms in Cambodia: Comparative Insights from the ASEAN, EU, Japan, and New Zealand  
*Leavsovath In and Phyrobish Ream*

**26**

Operationalizing ASEAN's AI Governance Guidelines: A Policy Implementation Study in Cambodia  
*Catherine Setiawan*



# Towards Ethical AI Governance on Digital Communication Platforms in Cambodia: Comparative Insights from the ASEAN, EU, Japan, and New Zealand



*In Leavsovath and Ream Phyrobish*

## I. Introduction

Artificial Intelligence (AI) is no longer regarded as a purely theoretical idea; it has been integrated into our daily lives. Its applications vary across domains, such as assisting in language translation, generating information, and reshaping human intelligence for problem-solving, decision-making, and continuous learning.<sup>1</sup> In the 21st century, AI is integrated into nearly every societal facet, including academia, development projects, military capabilities, and digital communication platforms to improve global connectivity.

However, AI brings both significant benefits and serious risks, depending on how it is implemented, used, and governed.<sup>2</sup> Digital communication platforms such as social media help people overcome barriers of time and distance and make information globally accessible, while AI integration provides powerful tools—like search and recommendation systems—that guide users through this vast flow of content.<sup>3</sup> On the other hand, people manipulate conversation by using Generative AI to spread disinformation or create harmful content that interrupts the flow of social networks.<sup>4</sup> Moreover, AI in detection systems or communication guidelines that serve as one ethical use of tracking down the AI-manipulated contents, somehow also become a form of weapon that people could abuse in mass reporting or harassment that leads to misinterpretation and punishment of innocent accounts.<sup>5</sup> Not

only does it affect the normal citizens, but it also brings a negative impact on government information and communication channels around the world, leading to the intervention with new regulations, policies, frameworks, or guidelines for ethical AI use.

In the context of Cambodia, fake news has posed a direct threat to national security and given a false alarm on sensitive topics to the public. The Ministry of Information of Cambodia has its own Fake News Monitoring Committee to track daily fabricated news. However, the focus leans more towards issue settlement rather than a preventive mechanism regarding AI-related misuse.<sup>6</sup> Therefore, the government collaborated with UNESCO to apply the AI Readiness Assessment Methodology in Cambodia and publish a national AI readiness report. This tool is the first systematic global framework designed to assess the country's readiness for ethical AI, examining five key areas: legal framework, social and cultural context, scientific and educational capacity, economic environment, and technical infrastructure.<sup>7</sup> Cambodia can strengthen its approach by learning from international organizations and developed countries that have already developed their AI policies, especially in the digital communication sector.

To develop an appropriate and reliable AI governance framework, certain international organizations and countries' AI ethics governance frameworks were proposed. This

chapter compares AI governance approaches from ASEAN, the European Union (EU), Japan, and New Zealand to identify lessons for Cambodia. ASEAN is included because Cambodia is part of the region and must align with regional digital policies. The EU provides strong legal frameworks like the General Data Protection Regulation (GDPR) and the AI Act that influence global digital governance. Japan is a key development partner supporting Cambodia's digital sector through technology and cooperation. New Zealand offers a practical model for ethical and transparent AI governance suitable for smaller countries. This study focuses on AI risks in the digital communication sector and uses these cases to propose policy options for Cambodia's ethical AI governance framework.

## II. AI Risk in Digital Communication Platform

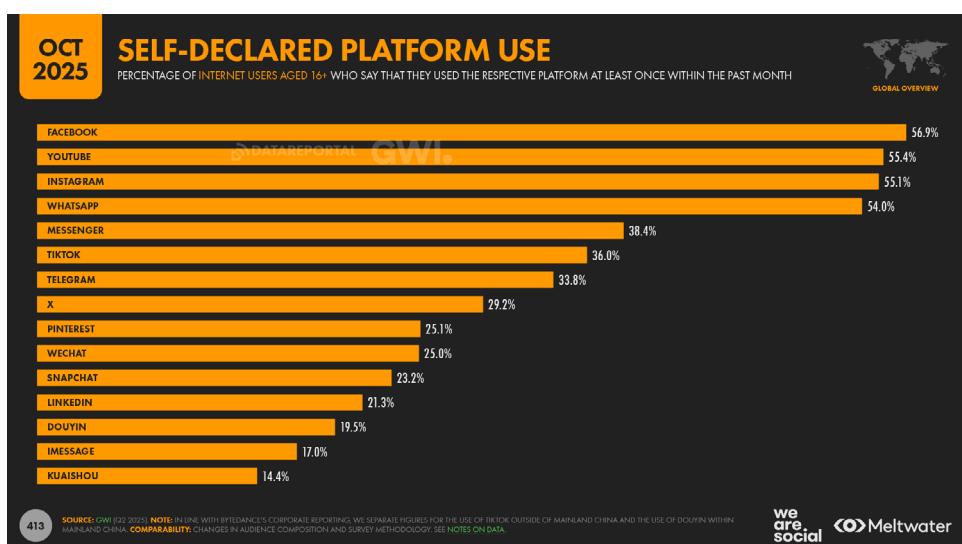
Regarding digital communication platforms, social media would serve as a bridge in connecting people both domestically and internationally. It has transformed the way people communicate by providing services for individuals to create accounts and

channels. Contemporary, different social media platforms are operated by different companies, each has distinct systems, principles, and ethical frameworks.

Based on the figure below, Facebook accounts for around 56.9% of global internet users, followed by YouTube and Instagram, according to GWI 2025. It is crucial to know why people engage in using social media. Brandtzaeg and Jan Heim 2009 found that people want to find new relations, make friends, and most importantly, access information.<sup>8</sup> Connecting to world contents, the Reuters Institute for the Study of Journalism (RISJ) found that over 37% of adults utilize platforms like Facebook to access and disseminate news.<sup>9</sup> This proves how user interaction and accessing information in digital communication platforms have slowly intertwined with misinformation.

People are using generative AI to create convincing misinformation and deepfakes, manipulating political narratives or voter behavior, and endangering country stability. For example, certain deep fakes on TikTok have polluted the United States election in 2024. Even the regional hegemon, like the US, is also prone to AI-manipulated content.<sup>10</sup>

Figure 1: Global Percentage of Self-Declared Platform Social Media Use



Source: GWI Digital Report cited from DATAREPORTAL

In most cases, social media operators respond by integrating with AI, particularly concerning their operational systems and content management. Developers would inject Natural Language Processing (NLP) to comprehend multilingual text, classifying content, and detecting public sentiment within posts.<sup>11</sup> AI algorithms are also on social media platforms to analyze user behavior and optimize user experience, by enhancing search functionality, but Emerging AI systems, including AI detection systems, are also being integrated into community guidelines to evaluate and manage inappropriate content or disinformation.<sup>12</sup>

In spite of that, AI systems still can produce unsafe or malicious outputs, yet transparency from developers remains limited. People unethically abuse the social media guidelines in place for personal gain. Therefore, even if private sector innovation accelerates, public safeguards and strong ethical governance from the government are mandatory to ensure protection and ethical use of AI without halting the innovation.<sup>13</sup>

### III. Cambodia's AI Readiness, Challenges, and Governance Gap

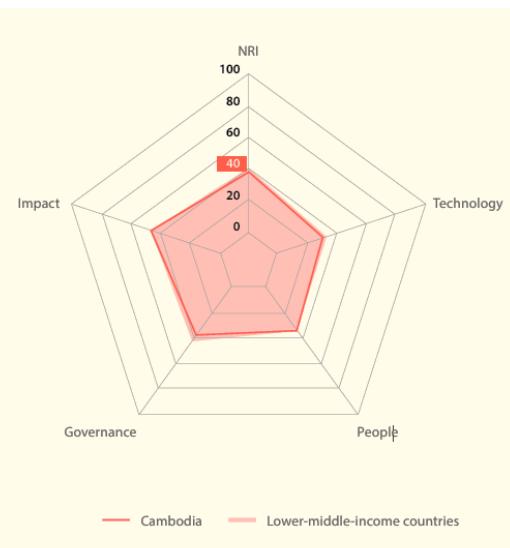
Since AI has become trendy, Cambodia strives its best to catch up with the global standard to utilize the benefits while managing the risks.

Cambodia's current position on AI readiness and its regulatory framework indicates significant room for growth. The Network Readiness Index 2024 ranks Cambodia 110th out of 133 countries, with a score of 35.65. Specific indicators highlight these gaps: in the technology pillar, "AI scientific publications" scored only 2.84 out of 100; in the people pillar, "AI talent concentration" and "Number of venture capital deals invested in AI" had no available index; and under the governance pillar, "Regulation of Emerging Technology" scored 35.66 out of 100.<sup>14</sup>

Moreover, despite these advancements, significant challenges persist. Cambodia's AI readiness is still developing, making it challenging to regulate every aspect, particularly using AI to combat disinformation and ensure ethical governance of media

Figure 2: Cambodia Network Readiness Index 2024

Network Readiness Index	Rank	Score (Out of 133)
	110	35.65
Pillar/sub-pillar	Rank	Score
<b>A. Technology pillar</b>	109	27.96
1st sub-pillar: Access	106	41.70
2nd sub-pillar: Content	89	19.19
3rd sub-pillar: Future Technologies	106	22.99
<b>B. People pillar</b>	100	33.64
1st sub-pillar: Individuals	41	54.32
2nd sub-pillar: Businesses	97	29.08
3rd sub-pillar: Governments	123	17.51
<b>C. Governance pillar</b>	121	36.60
1st sub-pillar: Trust	120	18.57
2nd sub-pillar: Regulation	117	48.45
3rd sub-pillar: Inclusion	110	42.79
<b>D. Impact pillar</b>	108	44.38
1st sub-pillar: Economy	100	24.91
2nd sub-pillar: Quality of Life	88	59.47
3rd sub-pillar: SDG Contribution	117	48.77



Source: Portulan Institute

outlets. For example, there have been serious scams using AI-generated images of a child who had lost their parents to ask for donations illegally.<sup>15</sup> Another example, deepfake videos and voice generators have been used to spread false information about Cambodian leaders or to gain political traction.<sup>16</sup> In addition, certain AI or Deepfakes were misused on women's privacy for social media attraction, which violated women's rights.<sup>17</sup> Those incidents were tackled using AI or modern technology by the Ministry of Information and the Ministry of Interior, but there should be a more safeguard mechanism, improved national AI expertise to ensure ethical and secure AI use in Cambodia.

The Royal government strives their best to introduce AI in "Digital Economy and Society 2021-2035". Following "Digital Government Policy 2022-2035", which outlines sub-strategies for AI adoption to support data governance, public service delivery, and economic innovation. As mentioned in the current government effort, Cambodia country readiness report has been published to evaluate the country's AI readiness. Moreover, the government has also announced progress on the draft of the National AI Strategy, which is going to be published soon. AI ethics were mentioned as a keyword for the assessment as a foundational principle within these policy discussions, laying the groundwork for responsible and accountable AI use. However, Cambodia's AI development is still in its early stages, and ethical governance will be very important after the assessments as learning from the existing policies of different countries and organizations is very significant.

## IV. AI Governance and Ethics Frameworks Review

### 1. EU and ASEAN as Organizational Frameworks

Both organizational approaches might look different, but they aim to make use of ethical principles in governing AI and provide protection to society.

The EU has anchored its AI strategy in a Human-centered legal foundation, aiming to make sure technology serves people first. This goes beyond simple ethics; the EU actively designs algorithms to include legal and economic checks against any violations of human rights, personal data, or privacy.<sup>18</sup> Crucially, the EU's Digital Education Action Plan 2021-2027 was adopted before the recent wave of generative AI tools and focuses more broadly on digital skills, media and information literacy, and strengthening educators' capacity. Its emphasis on critical thinking and responsible use of digital technologies nevertheless provides a useful basis for addressing emerging challenges such as generative-AI-driven misinformation.

On the other hand, ASEAN prefers a pragmatic approach, like "soft law" over strict regulation. Its key guides, including the *ASEAN Guide on AI Governance and Ethics* (2024),<sup>19</sup> with an expanded version on *Generative-AI* in (2025), with the non-binding principles from fairness to accountability.<sup>20</sup> The framework is flexible for ASEAN members with different AI readiness, and promotes interoperability and digital integration across the region.<sup>21</sup>

### 2. Japan and New Zealand as National Framework Case Studies

Organizational frameworks offer strong credibility; however, national frameworks from high-tech countries also serve as a good model. Japan, known as a technological leader in Asia, has adopted a few AI frameworks even amid its stagnation in innovation. Japan's contemporary position itself in a global race in ethical principles to align with its society 5.0. Its AI Governance Strategy, initiated by the Ministry of Internal Affairs and Communications (MIC) and the Ministry of Economy, Trade and Industry (METI), is divided into two primary eras: the period before Gen AI and the subsequent era after Generative AI. Below is the list of AI frameworks and ethical development in Japan since 2015.<sup>22</sup>

Table 1: History of AI and Its Governance Strategy in Japan

Timeline	Strategy or Policy	Guidelines	Institutions
Initial Period (≈2015 -2017)	AI Technology Strategy	MIC: Draft AI R&D Principles MIC: Draft AI R&D Guidelines for International Discussion	AIST – Artificial Intelligence Research Center (AIRC) RIKEN – Center for Advanced Intelligence Project (AIP)
OECD AI Principles & first national strategy (2019)	AI Strategy 2019: AI for Everyone	OECD Council Recommendation on AI (OECD AI Principles) MIC: AI Utilization Guideline CAO: Social Principles of Human-centered AI	AI Strategy Implementation Council AI Japan R&D Network
From Principle to Practice (≈2020-2022)	AI Strategy 2022	METI: AI Governance in Japan Version 1.1 METI: Governance Guidelines for Implementations of AI Principles Version 1.1	METI Expert Group on How AI Principles Should be Implemented AI Strategy Implementation Council
<b>Fourth AI Boom (Generative AI)</b> <b>AI Strategy Council (AISC) was established in May 2023</b>			
After Gen AI: integrated soft-law framework (2024-2025)	Consolidation of earlier guidelines into AI Guidelines for Business	MIC/METI: AI Guidelines for Business Version 1.0 and Version 1.1	MIC & METI as co-issuers Japan AI Safety Institute (J-AISI) within IPA (Information-technology Promotion Agency) National Institute of Information and Communications Technology (NICT): GPAI Tokyo Expert Center (July 2024)

Toward an AI Act-type framework (2024–2025)	Tentative Interim Report  Bill on the Promotion of Research, Development and Use of AI-related Technologies	The Interim Report itself is not a “principle” document but a policy/legislative design paper; it builds on earlier soft-law (AI Guidelines for Business, Governance Guidelines, etc.).	AI Strategy Council  Cabinet / Cabinet Office (CSTI) as the submitting authority for the AI Bill
---	---	---	--

*Source: Center for Strategic & International Studies | Japan Chairs*

Based on Table 1, before the Gen AI period, Japan began drafting “AI R&D Guidelines” in 2017 to promote best practices in using AI for research and fostering societal harmony.<sup>23</sup> Moreover, publication was the “AI Utilization Guidelines” in 2019, which aimed to promote the safe and reliable use of AI systems and products with Human-centered values for both business and public services.<sup>24</sup> Following the introduction of Gen AI, Japan issued additional guidelines, such as the “AI Guidelines for Business” and even a Bill on “AI Act,” to strengthen previous R&D guidelines.<sup>25</sup> However, these guidelines, so far, do not fully capture the ethical governance of Gen AI within digital communication platforms, though they do address certain aspects like transparency and fairness in providing or using AI services.

As for New Zealand’s framework, the Public Service AI Framework focuses on a “light-touch, proportionate, and risk-based approach”<sup>26</sup> on existing legislation rather than enacting new AI law.<sup>27</sup> What feature in this country’s framework is the integration of the Treaty of Waitangi, which includes the perspective of the Māori indigenous group on AI development? The frameworks emphasize human-centeredness, transparency, and accountability to support the safeguarding of each legislation.<sup>28</sup>

Both countries have different approaches to implementing their AI guidelines, but they are strengthening their AI guidelines to become binding legislation.

## V. Relevance and Applicability to Cambodia’s Policy Context

Cambodia’s draft National AI strategy is in the process of collecting more input from relevant stakeholders, with a specific focus on ethical governance under the Government plan. Hence, organizational frameworks from ASEAN and the EU, and national frameworks from Japan and New Zealand would support the finalization of the strategy, or further develop a specific focus on Digital Communication Platform governance to ensure its alignment with international.

### 1. Risk-Based Approach

The approach to risk management in AI governance is implicitly integrated into its non-binding framework, where principles like “security and safety” and “robustness and reliability” are mentioned.

For the EU, Cambodia could adopt AI ethics governance in education by using the EU’s Digital Education Action Plan (2021-2027). By adjusting the EU’s data rules for schools, Cambodia can create plans for teachers to use AI responsibly, which will encourage critical thinking.<sup>29</sup> A risk-focused approach that combines fact-checking and user-friendly tools for media awareness is needed to counter disinformation and improve digital skills related to AI ethics in Cambodia.

For ASEAN, ASEAN's 2025 Expanded Guide helps member states address AI risks. Though not required, Cambodia, as a member, might adopt the guide on AI governance and ethics. This guide focuses on generative AI and advises on handling deepfakes, misinformation, and copyright issues.<sup>30</sup> This decentralized model relies on individual nations to implement risk assessments tailored to their specific contexts and capacities.<sup>31</sup>

For Japan, Cambodia can learn from the coming AI Bill, which focuses on AI tech research, growth, and use. The bill's focus on risk-based governance offers a way to include legal and ethical rules with ways to check for violations like copyright issues, fake news, and deepfakes.<sup>32</sup>

For New Zealand, Light-touch, proportionate, and risk-based approach helps leverage existing legislation as "guardrails" (e.g., "safety and security") to guide public sector agencies in applying robust risk management, ensuring traceability of data, and minimizing harm from AI systems, which might be a viable option for Cambodia.<sup>33</sup>

## 2. Agile Regulation

Agile regulation is essential for maintaining pace with the rapid advancements in AI technology. Drawing on the European Union, this organizational AI Act provides for voluntary regulatory sandboxes, allowing high-risk AI systems to be tested under supervision before deployment.<sup>34</sup> The Act is risk-based and largely technology-neutral, while still defining special rules for some AI categories. Thus, Cambodia could draw on this approach for its own AI ethics governance on new AI products or media platforms related.

ASEAN's "soft law" approach could provide a degree of adaptability, allowing for updates like the 2025 Generative AI guide to address new technological risks without requiring complex legislation across ten diverse countries.<sup>35</sup> This enables rapid policy responses to emerging AI

challenges, advancing flexibility and regional interoperability for Cambodia as a member state.

From Japan, Cambodia can learn from the "AI Guidelines for Business", which focus on adaptive governance, with generative AI added as one of the agendas. Generative AI's flexibility helps with business tasks like communication, but its governance needs to be overseen without adding too much burden.<sup>36</sup> Cambodia could adopt this approach to see its credibility.

From New Zealand, a choice to rely on existing laws and principles-based frameworks as risks arise aims to prevent static governance from becoming obsolete, facilitating continuous monitoring and evolution of policies. Employing such monitoring policies would provide a practical ethical standard for the Cambodian AI approach.

## 3. Human-Centricity

Human-centered approaches to AI are always shaped by cultural settings and by the ways societies choose to deploy and govern these systems. In the EU, the Trustworthy AI Guideline mainly highlights accountability, transparency, and unbiased use of AI to enhance the system and improve human welfare. In communication platforms, the guidelines mentioned the protection of user rights; however, Cambodia's capacity to monitor the harmful uses of AI is currently limited. Therefore, Cambodia should explore ways to measure the level of human-centredness within its AI applications.

In ASEAN, its Human-centered approach to AI mainly places particular emphasis on the cultural context of people from different communities. The guide explicitly mentioned "Human-centricity" as a core principle, attempting to make an AI system to respect human-centered values and pursue benefits that are important for society as a priority. This principle is often interpreted through the lens of maximizing public good and ensuring AI serves diverse populations across the region.

Regarding Japan, its “Social Principles of Human-centered AI,” adopted in 2019, is somewhat similar to the other two organizational frameworks. These principles aim to promote fundamental human rights without affecting individuals’ rights, while also incorporating diversity and inclusivity in AI use.<sup>37</sup> Though Generative AI or AI in social media are not specifically named, ethical AI should still protect human rights. It is still important to have some rules for how people should act when using AI, mainly in places like social media, online news, and political communication, where AI-generated content can distort public debate and undermine trust. However, adopting a semi-governance approach to human behavior when using AI remains necessary, especially in a country like Cambodia, which is particularly vulnerable to misinformation.

Lastly in New Zealand, Public Service AI Framework also strongly promotes a Human-centered approach, emphasizing that AI should respect “the rule of law, democratic values and human rights and Labor rights” and ensure human oversight. A unique and highly significant aspect of New Zealand’s Human-centricity is the explicit incorporation of the Treaty of Waitangi and the critical consideration of *Te Ao Māori* (Māori worldviews), *mātauranga Māori* (Māori knowledge), and *tikanga* (Māori customs) in AI development.<sup>38</sup> This approach is also relevant for Cambodia, where AI can both promote cultural heritage and generate harmful side effects that require careful scrutiny.

In conclusion, the four AI governance approaches from the EU, ASEAN, Japan, and New Zealand help offer valuable insights for Cambodia’s development under the draft of the National AI Strategy. These can help improve the benchmark of the published RAM and further strengthen AI ethical governance adoption in the future.

## VI. Policy Synthesis and Actionable Recommendations

Cambodia’s draft National AI Strategy (2025-2030) would provide an excellent opportunity to leverage AI for national development while injecting necessary ethical safeguards.<sup>39</sup> The core principle of this strategy relies on a Human-centered approach that specifically respects Khmer culture, language, and societal values.<sup>40</sup> It is unfortunate that specific AI legislation remains absent, and foundational laws like the Personal Data Protection Law and cybersecurity frameworks are still in progress; therefore, strategic guidance is critically important.<sup>41</sup>

By synthesizing the ethical guidance policies, Cambodia’s policy direction is informed by both regional and global governance models, offering a pragmatic and flexible lens:

### Regional Alignment (ASEAN)

As an active ASEAN state member, Cambodia can directly embed the regional AI framework’s emphasis on transparency, fairness, and accountability. This foundational ethical guidance is crucial in combating the risks of misinformation and deepfakes within Cambodia’s growing media landscape.<sup>42</sup>

### Global Adaptation (EU, Japan, New Zealand)

Frameworks from international offer models for adaptive governance. For example, the EU’s approach supports AI system trials in regulatory sandboxes, while Japan focuses on the proposed AI Act to govern AI usage. New Zealand emphasizes robust checks within existing structures. These examples provide useful insights for flexible regulation that is less restrictive.<sup>43</sup>

## Cultural Preservation

A key Human-centered element is the commitment to ensure that AI development respects Khmer culture, language, and societal values, preventing misrepresentation by globally trained models. The MPTC is already proactively working to address this concern by building culturally aware AI models and multicultural benchmarking datasets.<sup>44</sup>

In synthesis, Cambodia can effectively combine these models, from adopting the overarching principle-based ASEAN guidance, to integrating pragmatic risk-based approaches from global partners, as well as firmly embedding its powerful example of culturally embedded Human-centricity into any sub-strategy of the draft strategy and future ethical governance development.<sup>45</sup>

## VII. Actionable Recommendations for the Royal Government of Cambodia

Cambodia needs to build explicit ethical guidelines for AI in digital communication, instead of only depending on the forthcoming National AI Strategy, considering AI's growing reach in these platforms.

### Adopt a Hybrid Regulatory Model for Ethical Definition

Cambodia should blend global principles with local adaptability by pursuing a dual-track approach:

- *Integrate Principle-Driven Frameworks:* Incorporate ASEAN's AI Governance and Ethics guides (2024), plus its 2025 work on generative AI, into relevant national policy. The focus should be on algorithmic bias and data misuse on communication platforms.
- *Adapt Existing Legal Structures:* Adapt current laws on cybersecurity, privacy, and digital rights to include AI uses on

social media. Doing so addresses the issue faster than creating new, separate laws and incorporates ethical standards into current regulations.

### Embed Ethical Oversight Institutionally

Cambodia must establish or utilize an existing mechanism to institutionalize ethical oversight, ensuring transparency and accountability:

- *Empower a Governance Body:* To advance the draft National AI Strategy, the government should either reinforce the current working group or form a specialized, multi-party team, preferably under the leadership of the MPTC.
- *Mandate and Focus:* This group would create ethical guidelines and tech standards emphasizing openness, fairness, and human control in AI. It would also supervise ethical risk reviews for digital tools like social media algorithms and represent national interests in ASEAN and global discussions, ensuring flexible responses to changes.

### Ensure Ethical Alignment in Legal Drafting

AI ethics should be part of the legal system to ensure AI is used responsibly in online media:

- *Legal Integration:* UNESCO's AI Readiness Report says AI ethics should guide the creation of current data protection and cybersecurity laws.
- *Rights Protection:* Ethical AI governance does not restrict freedom of expression; it aims to limit the harm caused by misinformation, algorithmic manipulation, and biased moderation while ensuring that digital rights are clearly defined and protected. International cooperation, such as UNESCO-led assessments, can help Cambodia develop rights-based and context-appropriate ethical standards for managing AI risks in digital media.

## **VIII. Conclusion**

In conclusion, the study thoroughly emphasizes that Cambodia has limited capability in governing AI ethical use and misinformation, which affects the livelihood of people and the operation of the government. Despite receiving support from UNESCO on the AI readiness assessment and already drafting the National AI Strategy, Cambodia needs to further strengthen ethical AI governance on digital communication platforms. The existing framework comparisons from the EU, ASEAN, Japan, and New Zealand would serve as benchmarks for designing context-appropriate policy approaches. Therefore, gradual development would help build a robust draft of the National AI Strategy and promote greater inclusiveness in ethical AI governance in Cambodian society.

## References

1. CSU Global. "What Is Artificial Intelligence (AI)?" Colorado State University Global, 2025. <https://csuglobal.edu/blog/how-does-ai-actually-work>.
2. Gow, Glenn. "AI's Double-Edged Sword: Managing Risks While Seizing Opportunities." *Forbes*, July 14, 2024. <https://www.forbes.com/sites/glenngow/2024/07/14/ais-double-edged-sword-managing-risks-while-seizing-opportunities/>.
3. Chen, Xinyan. "Effect Analysis of Digital Communication of Social Media from the Perspective of New Media." Paper presented at the 4th International Conference on Educational Innovation and Philosophical Inquiries, 2023. <https://doi.org/10.54254/2753-7064/10/20231242>.
4. Petropoulos, Georgios. "The Dark Side of Artificial Intelligence: Manipulation of Human Behaviour." *Bruegel*, 2022. <https://www.bruegel.org/blog-post/dark-side-artificial-intelligence-manipulation-human-behaviour>.
5. Meisner, Colten. "The Weaponization of Platform Governance: Mass Reporting and Algorithmic Punishments in the Creator Economy." *MediaWell, Social Science Research Council*, 2023. <https://mediawell.ssrc.org/citations/the-weaponization-of-platform-governance-mass-reporting-and-algorithmic-punishments-in-the-creator-economy/>.
6. Seangly, Phak. "Minister Warns of Daily Fake News Threat Undermining Cambodia's Stability." *Asia News Network*, 2024. <https://asianews.network/minister-warns-of-daily-fake-news-threats-undermining-cambodias-stability/>.
7. UNESCO. "Initiating AI Readiness Assessment: Shaping Cambodia's Future Technology." UNESCO, 2024. <https://www.unesco.org/en/articles/initiating-ai-readiness-assessment-shaping-cambodias-future-technology>.
8. Brandtzæg, Petter Bae, and Jan Heim. "Why People Use Social Networking Sites." In *Online Communities and Social Computing, Lecture Notes in Computer Science*. Berlin: Springer, 2009. [https://doi.org/10.1007/978-3-642-02774-1\\_16](https://doi.org/10.1007/978-3-642-02774-1_16).
9. Newman, Nic. "Overview and Key Findings of the 2025 Digital News Report." *Reuters Institute for the Study of Journalism*, 2025. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025/dnr-executive-summary>.
10. Bong, Shannon. "How AI Deepfakes Polluted Elections in 2024." *NPR*, December 21, 2024. <https://www.npr.org/2024/12/21/nx-s1-5220301/deepfakes-memes-artificial-intelligence-elections>.
11. Mohammad, Norazlinda, Norena Abd Karim Zamri, Mastura Roni, Siti Nur Izyandiyana Ab Hadi, Siti Fairuz Nurr Sadikan, and Sulaiman Mahzan. "Navigating AI Ethics in Malaysian Universities: Addressing Privacy, Integrity, and Bias." *International Journal of Research and Innovation in Social Science* 9 (2025). <https://dx.doi.org/10.47772/IJRRISS.2025.9010197>.
12. Salah, Mohammed, Hussam Al Halbusi, Fadi Abdelfattah, "May the force of text data analysis be with you: Unleashing the power of generative AI for social psychology research." *Computers in Human Behavior Artificial Humans* 1(7), 2023, DOI:10.1016/j.chbah.2023.100006.
13. Prilliadi, Hanif, "Harnessing AI for ASEAN's Future: Governance, Adoption, and Sustainability under the DEFA." *ERIA Policy Brief*, no. 2025-1, <https://www.eria.org/>

research/harnessing-ai-for-asean-s-future-governance--adoption--and-sustainability-under-the-defa.

14. Portulans Institute, *Network Readiness Index 2024: Accelerating Digital Transformation*. Accessed August 4, 2025, <https://download.networkreadinessindex.org/reports/data/2024/nri-2024.pdf>.
15. Ousa, Rin, "Rise in AI-Generated Scams Requires Quick Action." *Cambodianess*, 2025, <https://cambodianess.com/article/rise-in-ai-generated-scams-requires-quick-action>.
16. Rinith, Taing, "MPTC Condemns Deepfake Video of Senate President". *KhmerTimes*, 2025, <https://www.khmertimeskh.com/501709604/mptc-condemns-deepfake-video-of-senate-president/>.
17. Molika, Meas, "Deepfake Porn Targets Cambodian Women, Calls Grow for Data Protection Law." *Kiripost*, 2025, <https://kiripost.com/stories/cambodia-deepfake-porn-targets-cambodian-Women-calls-grow-data-protection-law>.
18. European Parliamentary Research Service, "EU Human-centered Approach to Artificial Intelligence". *European Parliament Briefing*, 2019, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRI\\_BRI\(2019\)640163\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRI_BRI(2019)640163_EN.pdf).
19. ASEAN, *ASEAN Guide on AI Governance and Ethics*, 2024, [https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics\\_beautified\\_20223\\_v2.pdf](https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_20223_v2.pdf).
20. ASEAN, *Expanded ASEAN Guide on AI Governance and Ethics – Generative AI*, 2025, <https://asean.org/wp-content/uploads/2025/01/Expanded-ASEAN-Guide-on-AI-Governance-and-Ethics-Generative-AI.pdf>.
21. *Ibid.*
22. Greene, Samantha Hoffman, and Akira Igata. *Norms in New Technological Domains: Japan's AI Governance Strategy*. Washington, DC: Center for Strategic and International Studies (CSIS), July 20, 2021. <https://www.csis.org/analysis/norms-new-technological-domains-japans-ai-Governance-strategy>.
23. Conference Toward AI Network Society, "Draft AI R&D Guidelines for International Discussion." Accessed July 21, 2023. [https://www.soumu.go.jp/main\\_content/000507517.pdf](https://www.soumu.go.jp/main_content/000507517.pdf).
24. Ministry of Internal Affairs and Communications (Japan), *AI Utilization Principles* (Tokyo: Government of Japan, 2019), [https://www.soumu.go.jp/main\\_content/000658284.pdf](https://www.soumu.go.jp/main_content/000658284.pdf).
25. *Ibid.*
26. Valentine, Nick, Emma Moran, M. Fitzpatrick-Cockram, and Vivian Mitchell, "New Zealand's Public Service AI Framework: Guiding Responsible Innovation." *DLA Piper*, 2025, <https://www.dlapiper.com/insights/publications/2025/02/new-zealands-public-service-ai-framework-guiding-responsible-innovation>.
27. Ministry of Business, Innovation & Employment, *New Zealand's Strategy for Artificial Intelligence: Investing with Confidence*, 2025, <https://www.mbie.govt.nz/dmsdocument/30888-new-zealands-strategy-for-artificial-intelligence-investing-with-confidence>.
28. Digital.govt.nz, *Public Service AI Framework*, 2025, <https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/artificial-intelligence/public-service-artificial-intelligence-framework>.
29. *Ibid.*
30. *Ibid.*
31. *Ibid.*

32. *The Peninsula Qatar, "Japan's Cabinet Approves AI Bill to Balance Risk Management, R&D."* *The Peninsula Qatar, March 1, 2025.* <https://thepeninsulaqatar.com/article/01/03/2025/japans-cabinet-approves-ai-bill-to-balance-risk-management-rd>.
33. *Ibid.*
34. *European Parliament, "EU AI Act: First Regulation on Artificial Intelligence."* *European Parliament, June 1, 2023.* <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence#ai-regulation-in-europe-the-first-comprehensive-framework-4>.
35. *Ibid.*
36. *Ibid.*
37. *Albagli, David, Toshio Dokei, Tim Hickman, and Arthur Mitchell, Regulation of Artificial Intelligence in Europe and Japan.* *Jdsupra.* August 25, 2020. <https://www.jdsupra.com/legalnews/regulation-of-artificial-intelligence-53854/>.
38. *Ibid.*
39. *UNESCO, Recommendation on the Ethics of Artificial Intelligence, 2023,* <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.
40. *Ministry of Post and Telecommunication, Draft National Artificial Intelligence Strategy 2025-2030, 2025,* <https://data.opendvelopmentcambodia.net/en/library-record/draft-National-artificial-intelligence-strategy-2025-2030-version-5>.
41. *Hongseng, Rov, "From Reality to Policy: What Cambodia's AI Readiness Assessment Report Revealed."* *Kiripost, 2025,* <https://kiripost.com/stories/from-reality-to-policy-what-cambodias-ai-readiness-assessment-report-revealed>.
42. *Cambodgemag.com, "The Digital Awakening: How Artificial Intelligence Is Reshaping Cambodia's Media Landscape."* *Cambodgemag, 2025,* <https://www.cambodgemag.com/en/post/The-digital-awakening-how-artificial-intelligence-is-reshaping-cambodia-s-media-landscape>.
43. *AI Cambodia, Ensuring Cambodian Identity in the Age of AI.* *June 23, 2025.* <https://ai.gov.kh/ensuring-cambodian-identity-in-the-age-of-ai/>.
44. *Ibid.*
45. *Ibid.*



Mr. In Leavsovath

*In Leavsovath is currently pursuing a master's degree in economics at the Graduate School of International Cooperation Studies, Kobe University. Prior to his academic journey in Japan, he worked as an Associate Research Assistant at the Cambodia Development Resource Institute (CDRI) and later as a Policy Researcher with the Digital Economy and Business Committee (DEBC) under the Ministry of Economy and Finance. He recently joined a youth policy competition on "AI for Peace" at Fudan University in Shanghai. His research has primarily focused on digital education, and he has also contributed to policy briefs related to emerging technologies.*



Mr. Ream  
Phyrobish

*Ream Phyrobish is currently a Public Policy Analyst at the Ministry of Economy and Finance | Non-Bank Financial Services Authority. He has recently completed a Master of Public Policy from Victoria University of Wellington, New Zealand, as a recipient of the 2024 Maanaki New Zealand Scholarship. His professional work currently directly contributes to the ongoing development of the International Financial Center of Cambodia (IFCC). During his studies, he contributed to a range of policy analyses, offering options and recommendations on topics including public resource management, algorithmic bias in AI-driven diagnostics, economic and health policy, and risk and cost-benefit analysis on development projects. Prior to his academic journey, he gained experience working with NGOs, development agencies, and the private sector. His professional background focuses on trade facilitation regulations.*

# Operationalizing ASEAN's AI Governance Guidelines: A Policy Implementation Study in Cambodia



Catherine Setiawan

## I. Introduction

Advances in generative artificial intelligence (AI) have outpaced regulatory frameworks in Southeast Asia, creating a mismatch between rapid adoption and the establishment of necessary governance.<sup>1</sup> This lag has created a regional governance gap, a risk that could undermine public trust, safety, and cross-border collaboration on AI deployment. As a response, ASEAN member states released the *ASEAN Guide on AI Governance and Ethics* (ASEAN AI Guide 2024) in February 2024, outlining shared principles and recommended best practices for implementing trustworthy AI across the region.<sup>2</sup> As a landmark policy document, the guide aims to build consumer trust and enable the responsible, cross-border deployment of AI services and solutions. This guide has since been expanded with the release of the *Expanded ASEAN Guide on AI Governance and Ethics-Generative AI* (Gen AI Guide 2025). The Gen AI Guide complements the ASEAN AI Guide 2024 by addressing policy considerations specific to generative AI. It explores both the opportunities and risks associated with Gen AI and recommends a set of policy actions to support its responsible adoption across ASEAN.

While the ASEAN AI Guides are a critical step toward regional alignment, actual implementation across member states has been uneven due to differences in institutional capacity, digital infrastructure, legal systems, and political priorities. Against this backdrop,

this chapter examines Cambodia in terms of its progress in operationalizing the ASEAN AI Guide guidelines into national policy and regulatory frameworks. This chapter will also recommend a practical, scalable roadmap for harmonizing AI governance implementation in Cambodia, tailored to national capacities. This chapter has five sections, starting from the introduction and background, an overview of ASEAN AI Governance and its progress in Cambodia, before providing conclusions and recommendations to improve the Cambodia's AI governance agenda for its AI Governance Strategy.

## II. Background

AI governance is critically important in Cambodia as the country accelerates its digital transformation and seeks to harness artificial intelligence for economic growth, public service delivery, and innovation. With rapidly emerging technologies, there is a pressing need to ensure that AI systems are developed and used responsibly, ethically, and safely. Effective governance helps prevent harms such as bias, misinformation, and privacy violations, while fostering public trust and accountability. It also enables the country to align with regional and international standards, particularly the ASEAN AI Governance and Ethics Guidelines, supporting cross-border cooperation and digital trade. Moreover, strong governance ensures that AI benefits are distributed fairly, helping to bridge digital divides and protect the rights and well-being of all Cambodians.

### III. ASEAN AI Governance

AI governance refers to the processes, standards, and guardrails that help ensure AI systems and tools are safe and ethical.<sup>3</sup> Its frameworks guide the research, development, and use of AI to promote safety, fairness, and respect for human rights. Effective AI governance involves oversight mechanisms to manage risks like bias, privacy violations, and misuse while encouraging innovation and trust. It offers a structured framework to reduce these risks, incorporating policies, regulations, and data governance. Regulation of AI is needed to address risks to vulnerable populations, protect privacy, and address ethical issues of AI use.<sup>4</sup> The ASEAN AI Governance chronology, development, and comparison can be seen in Table 1 below.

#### 1. ASEAN AI Guide 2024

The ASEAN AI Guide 2024,<sup>5</sup> was endorsed at the 4th ASEAN Digital Ministers' Meeting in February 2024. This guide is the first regional AI ethics framework that offers common principles and best practices for AI deployment in commercial and non-military contexts. It provides a flexible, risk-based approach tailored to the diverse digital readiness and governance capacities of ASEAN member states. This guide consists of seven key principles:

- Transparency and Explainability: AI systems should be designed and operated in a way that allows people to understand how they work and make decisions.
- Fairness and Equity: AI systems should be unbiased and avoid discrimination against any individual or group.
- Security and Safety: AI systems must be secure and resilient throughout their lifecycle.
- Robustness and Reliability: AI systems should perform consistently and as intended.
- Human-centricity: AI should prioritize human well-being and autonomy.
- Privacy and Data Governance: AI systems must uphold data privacy and ensure responsible data management.
- Accountability and Integrity: There should be clear responsibility for the development, deployment, and use of AI systems, with mechanisms in place to address any harms they may cause.

*Table 1. ASEAN AI Governance: Chronology, Development and Comparison*

Document	Date	Focus Area	Purpose
ASEAN AI Guide 2024	February 2024	Ethical principles	Establish regional baseline for AI
Gen AI Guide 2025	January 2025	Generative AI risks/benefits	Address Gen AI-specific policies
ASEAN Responsible AI Roadmap (2025–2030)	July 2025	Implementation strategy	Operationalize guides across ASEAN

## 2. Gen AI Guide 2025

The expanded Gen AI Guide 2025,<sup>6</sup> was released at the 5th ASEAN Digital Ministers' Meeting in January 2025. This version extends the original ASEAN AI principles to address the unique risks and opportunities of generative AI technologies, such as mistakes and anthropomorphism, inaccurate responses and disinformation, deepfakes, impersonation, fraudulent and malicious activities, infringement of intellectual property rights, privacy and confidentiality, and propagation of embedded biases. This Guide also explores frontier and systemic risks posed by long-term evolution in the capabilities of highly advanced Gen AI models, but which are not widespread in ASEAN at this time. The guide introduces additional considerations for the seven guiding principles set out in the ASEAN AI Guide 2024, namely:

- Need for transparency around content origin and model behavior; disclosure when outputs may be synthetic or hallucinated
- Awareness of bias propagation in generated content; ensuring linguistic, cultural, and demographic inclusivity in training data
- Mitigation of misuse risks: deepfakes, impersonation, prompt injection, data

poisoning; adversarial vulnerabilities

- Addressing hallucinations and factual inaccuracies, establishing incident-reporting protocols and risk assessment regimes
- Introducing shared responsibility across developers, deployers, and users; human oversight over content-generated decisions
- Strengthening IP compliance, anonymization, and traceability; preventing reconstruction of private data in outputs
- Clarifying liability among ecosystem actors; requiring governance frameworks for lifecycle accountability

Moreover, this guide also offers a range of policy recommendations for ASEAN to consider for promoting the trusted and responsible use of Gen AI in the region while addressing the risks previously mentioned. Recommendations are especially related to accountability, data, trusted development and deployment, incident reporting, testing and assurance, security, content provenance, safety and alignment research & development, and AI for public good. Details can be seen in Table 2 below.

Table 2. Key Policy Recommendations for Generative AI use in ASEAN

Explanation
<p><b>A. Establish Shared Accountability Across the Ecosystem</b></p> <ul style="list-style-type: none"> <li>• Clarify roles and responsibilities among AI developers, deployers, distributors, and end users.</li> <li>• Promote responsible behavior and decision-making throughout the AI lifecycle.</li> </ul>
<p><b>B. Ensure High-Quality, Lawful, and Diverse Training Data</b></p> <ul style="list-style-type: none"> <li>• Use datasets that are accurate, representative, legally obtained, and inclusive (especially of ASEAN languages and cultures).</li> <li>• Protect against embedded bias and discrimination in training data.</li> </ul>

### C. Design for Trustworthy Development and Deployment

- Integrate safety-by-design, explainability, and alignment mechanisms into model architecture.
- Support “red teaming” and adversarial testing before deployment.

### D. Implement Incident Reporting Mechanisms

- Create clear protocols to report, document, and respond to AI system failures or misuse (e.g., deepfake abuse or hallucinations).
- Promote regional information-sharing across ASEAN.

### E. Develop Testing, Assurance, and Evaluation Standards

- Use evaluation benchmarks and risk assessments tailored to ASEAN contexts (language, use cases, digital maturity).
- Align with international standards while allowing local flexibility.

### F. Strengthen AI Security

- Protect Gen AI systems from vulnerabilities such as prompt injection, adversarial attacks, and data leakage.
- Encourage “security by design” throughout model development and deployment.

### G. Enable Content Provenance and Traceability

- Promote watermarking, disclosure labels, and other tools to verify whether content was AI-generated.
- Support user awareness and content authentication infrastructure.

### H. Invest in Safety and Alignment R&D

- Fund collaborative regional research on value alignment, interpretability, and red-teaming methods.
- Create testbeds and sandboxes across ASEAN Member States.

### I. Leverage Gen AI for Public Good

- Encourage applications in education, healthcare, disaster response, and other sustainable development goals (SDGs).
- Promote access for underserved communities and small businesses.

Source: *Gen AI Guide 2025*

### 3. ASEAN Roadmap on AI 2025-2030

The ASEAN Responsible AI Roadmap (2025-2030),<sup>7</sup> is developed to provide actionable steps for ASEAN policymakers and stakeholders to create the ideal conditions for responsible AI to flourish in the region, as well as for ASEAN Member States (AMS) to leverage and enable responsible AI in a meaningful, impactful, and sustainable manner by 2030. This roadmap is developed to complement the ASEAN AI Guide 2024 and Gen AI Guide 2025. The roadmap is divided into "Cross-Cutting and Targeted" activities. Cross-Cutting activities were designed for ASEAN and its sectoral bodies/committees to coordinate efforts, approaches, and resources for integrated AI operationalization; while Targeted activities were designed to give AMS tailored objectives and outcomes for AI operationalization based on the maturity/readiness level of their national AI ecosystems. Details can be seen in Table 3 below.

*Table 3. Cross-Cutting and Targeted Activities: Pillars*

Activities	Pillars
<b>Cross Cutting</b>	<p>Skills &amp; Capacity Building: Improving the skills and knowledge that are essential for a workforce to benefit from a digitalized global economy, with an emphasis on AI-related know-how.</p> <p>Fairness &amp; Inclusion: Enhancing the visibility and voice of vulnerable, at-risk, or marginalized communities in the design and implementation of AI products, services, policies, and regulations.</p> <p>Governance &amp; Participation: Foster an inclusive, representative, and consultative approach in the development of AI governance and regulatory frameworks to ensure that AI systems remain Human-centered and inclusive.</p> <p>Integration &amp; Cooperation: Expanding the platforms for regional and global dialogue around key AI trends and issues.</p>
<b>Targeted</b>	<p>Internal Governance Structures and Measures: Promote awareness and support for the development and implementation of national AI policies, frameworks, strategies, and regulations.</p> <p>Skills &amp; Knowledge for Responsible, AI-Augmented Decision-Making: Promote and strengthen skills, capabilities, and knowledge-sharing among the public and private sectors, educational institutions, and civil society to develop and deploy AI responsibly.</p> <p>Risk Mitigation, Monitoring Mechanisms, and Operations Management: Strengthen and promote the use of risk mitigation and monitoring and evaluation tools across the public and private sectors to mitigate risks and biases, minimize harm from AI, and align with responsible AI principles.</p> <p>Stakeholder Coordination &amp; Regional Cooperation on AI: Promote regional efforts to develop common principles and a guiding approach to operationalizing responsible AI and align national efforts with the work of the ASEAN Working Group on AI Governance.</p>

*Source: ASEAN Roadmap on AI 2025-2030*

## IV. Cambodia AI Governance

Cambodia does not have a comprehensive AI law, but Sub-Decree 110 dated 21 July 2017 on ICT Licensing Regime indirectly affects AI suppliers, including importers, distributors, and developers of AI.<sup>8</sup> Thus far, Cambodia,<sup>9</sup> is making progress toward developing a National Artificial Intelligence Strategy (NAIS), with the vision to *"leverage artificial intelligence to increase productivity and generate added values, uplifting citizens' quality of life, strengthening and broadening the bases of sustainable and inclusive economic and social development, and building a prosperous future of Cambodia's digital technology"*. The aim is to improve citizens' literacy and skills, leverage AI for enhancing digital government, promote economic development, enhance the AI ecosystem, and manage AI risks.<sup>10</sup>

The Ministry of Post and Telecommunications (MPTC) is mandated to lead the work related to AI backbone infrastructure and AI governance and to coordinate with relevant ministries and institutions to develop policy frameworks and strategies on AI governance. While the Ministry of Industry, Science, Technology & Innovation (MISTI), is mandated to lead the work related to AI research and development and to coordinate with relevant ministries and institutions to prepare a roadmap on AI research and development.<sup>11</sup> The AI Strategy itself is developed by MPTC, in collaboration with United Nations Educational, Scientific and Cultural Organization (UNESCO) and United Nations Social and Economic Commission for Asia and the Pacific (UNESCAP) with the objective of strengthening Cambodia's AI policy and AI governance frameworks.<sup>12</sup> The country's AI preparedness was assessed using UNESCO's Readiness Assessment Methodology (RAM),<sup>13</sup> which provides a comprehensive evaluation of its capacity to govern AI systems ethically and inclusively. The MPTC and UNESCAP are collaborating on the formulation of an NAIS aligned with the Cambodia Digital Economy and Society Policy Framework 2021-2035 and the Digital Government Policy 2022-2035.<sup>14</sup>

On 11<sup>th</sup> June 2025, the MPTC made an announcement on the progress of preparing and opening public consultation on Draft National AI Strategy,<sup>15</sup> and inviting the public to provide feedback on the draft of the National AI Strategy (version 4), to regulate the fast-developing technology in a country with no standalone AI laws and limited data protection.<sup>16</sup> The Draft AI Strategy (Latest online: version 5) has six key priorities on human resource and development; data, infrastructure and key technologies; promoting AI adoption and development for Digital Government; Promoting sectoral adoption and development (education, healthcare, agriculture, manufacturing, tourism, MSMEs, finance, transport, digital); promoting ethical and responsible AI, promoting collaboration, research development and innovation.<sup>17</sup>

At the recent 2025 World Artificial Intelligence Conference and High-Level Meetings on Global AI Governance in Shanghai, China (26-28 July 2025),<sup>18</sup> H.E. Dr. Hul Seingheng (the Undersecretary of State at the MISTI) outlined Cambodia's commitment to building a future-ready digital economy by integrating AI into industry, SMEs, and all socio-economic sectors. The efforts focus on building full-bodied digital infrastructure, promoting AI integration, and nurturing a technology-friendly policy environment. He also emphasized Cambodia's priorities for global AI governance, including investment in human capital, regional fellowships, ethical AI use, and cross-border research collaboration. MPTC said the draft strategy will set ethical standards, transparency rules, and oversight mechanisms.

Table 4. Cambodia AI Governance Progress

Area	Status / Focus
Institutional Framework	MPTC is explicitly leading the planning and drafting of the national AI strategy 2025-2030
Ethical Readiness	UNESCO-supported AI RAM is completed, with the report already published; its recommendations feed directly into the draft national AI strategy 2025-2030
ASEAN Collaboration	Cambodia participates in ASEAN AI governance processes, such as the co-adoption of the ASEAN AI Guide (2024), the Expanded GenAI Guide and ASEAN Responsible AI Roadmap (2025), and hosts AI adoption side events and runs national consultations and translations of these documents.
Key Policy Pillars	Human resource and development; data, infrastructure, and key technologies; promoting AI adoption and development for Digital Government; Promoting sectoral adoption and development (education, healthcare, agriculture, manufacturing, tourism, MSMEs, finance, transport, digital); promoting ethical and responsible AI, promoting collaboration, research development, and innovation.
Timeline	<p>11 June 2025: MPTC opened a public consultation for the Draft National AI Strategy.</p> <p>1-2 July 2025: National workshop on the Draft NAIS and launch of the AI Readiness Assessment Methodology report</p> <p>2 July 2025: the latest publicly available text is the <i>Draft National Artificial Intelligence Strategy 2025-2030 (Version 5)</i></p> <p>August 2025: According to MPTC's April 2025 press release on AI governance, the plan was to submit the finalized NAIS to the National Digital Economy and Society Council for review and decision in August 2025. As of late 2025, publicly available sources still refer to the document as a <i>draft</i>, and there is no official announcement confirming its formal approval."</p>

Source: Author's compilation based on UNESCO (2025), MPTC (2025), ASEAN (2024)

As of mid-2025, Cambodia continues to build its legal and policy foundations for AI governance primarily through national strategies and frameworks, with binding legislation still in progress.<sup>19</sup> While the country has made strides in improving digital readiness,<sup>20</sup> several areas remain under development, including<sup>21</sup> enhancing cybersecurity capabilities, integrating AI technologies into public sector operations, and increasing investment in research, talent development, and infrastructure to support more advanced AI applications.

## V. Cambodia Integration of ASEAN AI Guidelines into National Governance Efforts

Despite the existing gaps, Cambodia remains actively engaged at the regional level, contributing to the ASEAN AI Guide 2024, the ASEAN Responsible AI Roadmap, as well as the Gen AI Guide 2025.<sup>22</sup> Cambodia engaged in ASEAN-led consultations, and its AI framework includes considerations consistent with ASEAN's seven guiding principles and Gen AI guide 2025, emphasizing on ethical standards, transparency, Human-centered AI, and regional collaboration, incorporating principles like transparency, fairness, accountability, data privacy, and human oversight into national planning core principles from the ASEAN frameworks.<sup>23</sup> Moreover, during the ASEAN COSTI-85 meeting in June 2024, Cambodia hosted a high-level policy dialogue involving ASEAN states, The Asia Foundation, and Stanford's Human-Centered AI Institute on fairness, bias, and human-centered AI adoption. However, full integration is ongoing, as the national strategy is still under finalization.

## VI. Conclusion

Cambodia is making progress toward developing a national governance framework for AI and is actively shaping its AI strategy in alignment with ASEAN's ethical and human-centered frameworks, though full integration remains in progress. Cambodia also demonstrates strong regional engagement and commitment to ASEAN's AI governance principles, despite its draft national AI strategy is still being finalized. Thus far, Cambodia's draft AI strategy reflects ASEAN values of ethics, transparency, and human oversight, reinforcing its role in regional AI governance. Despite remaining gaps, Cambodia is positioning itself as a committed regional partner in advancing responsible and inclusive AI governance.

## VII. Recommendation

Cambodia can better promote and harmonize AI Governance nationally, through:

- Improving AI Integration in the public sector: Cambodia can develop AI procurement guidelines and standards for public sector use, and train civil servants on AI literacy and management.
- Improving gaps in research, talent, and infrastructure: Cambodia needs to expand digital literacy programs nationwide, targeting marginalized and rural populations; establish scholarships, fellowships, and exchange programs focused on AI and data science, and upgrade digital infrastructure (broadband, cloud computing) to support AI experimentation and deployment. Additionally, funding for AI R&D in universities and innovation hubs needs to be increased.
- Improving digital readiness: private sectors and startups can be encouraged to develop localized AI solutions, as well as regional collaboration on shared AI tools and resources to leverage ASEAN economies of scale.

Cambodia also needs to develop its cybersecurity capabilities through partnering with ASEAN and international bodies for cybersecurity capacity building and knowledge sharing, to be able to invest in strengthening the national cybersecurity infrastructure with AI-specific safeguards.

## References

1. Isono, Ikumo, and Hilmy Prilliadi, *Accelerating Artificial Intelligence Discussions in ASEAN: Addressing Disparities, Challenges, and Regional Policy Imperatives*. ERIA Discussion Paper Series No. 488. Jakarta: Economic Research Institute for ASEAN and East Asia (ERIA), November 2023, <https://www.eria.org/uploads/media/discussion-papers/FY23/Accelerating-AI-Discussions-in-ASEAN-.pdf>.
2. *Ibid.*
3. IBM, "AI Governance." IBM Think, Accessed July 10, 2025, <https://www.ibm.com/think/topics/ai-governance>.
4. Yohanes, Widodo, "Challenges of Regulating Artificial Intelligence." *Kompas*, January 14, 2024, <https://www.kompas.id/baca/english/2024/01/14/en-tantangan-meregulasi-kecerdasan-buatan>. Accessed July 20, 2025.
5. Association of Southeast Asian Nations (ASEAN), *ASEAN Guide on AI Governance and Ethics*. February 1, 2024, <https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics.pdf>.
6. Association of Southeast Asian Nations (ASEAN), *Expanded ASEAN Guide on AI Governance and Ethics – Generative AI*, January 25, 2025, <https://asean.org/wp-content/uploads/2025/01/Expanded-ASEAN-Guide-on-AI-Governance-and-Ethics-Generative-AI.pdf>.
7. Association of Southeast Asian Nations (ASEAN), *ASEAN Responsible AI Roadmap (2025–2030)*, July 2025, <https://asean.org/wp-content/uploads/2025/07/ASEAN-Responsible-AI-Roadmap-2025-2030.pdf>.
8. Ministry of Posts and Telecommunications (Cambodia), *Draft National Artificial Intelligence Strategy 2025–2030, Version 5* (Phnom Penh: Ministry of Posts and Telecommunications, 2024), 3.
9. Chhem Kieth Rethy, Seingheng Hul, and Mounh Noy, *AI Landscape in Cambodia: Current Status and Future Trends* (Phnom Penh: Ministry of Industry, Science, Technology & Innovation (MISTI), May 2023), ISBN 978-9924-600-09-1.
10. Ministry of Post and Telecommunications, *Draft National AI Strategy*, 5.
11. Royal Government of Cambodia, Council of Ministers, Letter No. 660, April 21, 2025.
12. Ministry of Post and Telecommunications (Cambodia), "Cambodia Advances AI Ambitions with UNESCO Report, but Gaps Remain in Strategy, Cyber-Security and Data Governance," *Cambodia Investment Review*, July 3, 2025.
13. UNESCO, "Readiness Assessment Methodology: A Tool of the Recommendation on the Ethics of Artificial Intelligence." August 28, 2023.
14. United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP), "The Deep Dive Workshop for Cambodia's National AI Strategy," 7–8 April 2025, Phnom Penh, Cambodia, Accessed 22 October 2025,
15. Ministry of Post and Telecommunications (Cambodia), *Announcement on the Progress of Preparing and Opening Public Consultation on Draft National AI Strategy*, June 11, 2025, [https://opendevcambodia.net/pdf-viewer/?pdf=files\\_mf/1749711516mptc\\_en1.pdf](https://opendevcambodia.net/pdf-viewer/?pdf=files_mf/1749711516mptc_en1.pdf).
16. Seng Mengheng, "Cambodia Moves Cautiously towards AI Regulation amid Rising Concerns." *KiriPost*, June 23, 2025. Accessed 3 August 2025.
17. Ministry of Post and Telecommunications, *Draft National AI Strategy*, 5.

18. Ministry of Information (Cambodia), "Cambodia Is Developing National Artificial Intelligence Strategy to Ensure Safety, Ethics and Innovation," Ministry of Information, July 30, 2025, <https://www.information.gov.kh/articles/165606>.
19. United Nations Cambodia, "Workshop to Prepare Cambodia for Artificial Intelligence," UN Cambodia, accessed August 3, 2025, <https://cambodia.un.org/en/290551-workshop-prepare-cambodia-artificial-intelligence>.
20. UNESCO, "Initiating AI Readiness Assessment: Shaping Cambodia's Future Technology," UNESCO, accessed August 3, 2025, <https://www.unesco.org/en/articles/initiating-ai-readiness-assessment-shaping-cambodias-future-technology>.
21. Ministry of Post and Telecommunications (Cambodia), "Cambodia Advances AI Ambitions with UNESCO Report, but Gaps Remain in Strategy, Cyber-Security and Data Governance," Cambodia Investment Review, July 3, 2025.
22. Khmer Times, "Cambodia Hosts AI Adoption Side Event at ASEAN Meeting," Khmer Times, June 8, 2024, <https://www.khmertimeskh.com/501500237/cambodia-hosts-ai-adoption-side-event-at-asean-meeting>.
23. Ministry of Post and Telecommunications (Cambodia), "The ASEAN Guide on AI Governance and Ethics," [ai.gov.kh/the-asean-guide-on-ai-governance-and-ethics/](https://ai.gov.kh/the-asean-guide-on-ai-governance-and-ethics/).



**Ms. Catherine Setiawan**

*Catherine Setiawan holds a Master of Public Administration from Cornell University, USA, supported by a full scholarship from USAID Prestasi, and a Bachelor of Law degree from Pelita Harapan University, Indonesia. She currently serves as Country Coordinator for the Global Index on Responsible AI (GIRAI), leading assessments across ASEAN countries and contributing to regional benchmarking on responsible AI governance. Previously, Catherine was Project Coordinator for the International Telecommunication Union in Southeast Asia, leading regional AI governance initiatives to support social and economic development, and a Regulatory Reform Adviser under the Australia-Indonesia Partnership, supporting digital economy, data governance, and trade policy reforms. She likes to meditate and practice yoga during her free time.*



# STRATEGY FOR THE DEVELOPMENT OF AI

**38**

Readiness Assessment and the Development of National AI Strategy  
in Cambodia

*Pisal Chanty and Bunhong Taing*



# Readiness Assessment and the Development of National AI Strategy in Cambodia



Chanty Pisal and Bunhong Taing

## I. Introduction

Artificial Intelligence (AI) has rapidly moved to the forefront of national digital agendas worldwide. By 2025, at least 75 countries had adopted national AI strategies or formal frameworks, reflecting AI's growing role in economic policy and public-service reform.<sup>1 2</sup> In Cambodia, the government likewise views AI as a catalyst for digital transformation and sustainable development. The Royal Government's Pentagonal Strategy – Phase I (2023) elevated "Technology" as one of its five key priorities, underscoring digital innovation as fundamental to achieving Cambodia's Vision 2050.<sup>3</sup> Building on this vision, Cambodia enacted foundational digital policies – notably the **Digital Economy and Society Policy Framework 2021–2035** and the **Digital Government Policy 2022–2035** – which identify AI among the transformative tools to modernize infrastructure, spur public-sector innovation, develop human capital, and establish standards for the digital age.<sup>4 5</sup>

At the same time, Cambodia has committed to international and regional frameworks for ethical technology governance. In November 2021, it joined all UNESCO Member States in adopting the *Recommendation on the Ethics of Artificial Intelligence*, the world's first global AI ethics instrument.<sup>6</sup> This commitment, along with Cambodia's active role in ASEAN's efforts on AI ethics and governance, set the stage for a systematic assessment of the nation's AI readiness. By late 2024, Cambodia embarked on a comprehensive **AI Ethics Readiness**

**Assessment** using UNESCO's methodology (RAM) – an internationally recognized tool to evaluate a country's preparedness to govern AI in a responsible, human-centric manner.<sup>7</sup> The aim was not only to take stock of the current AI ecosystem, but also to generate evidence-based recommendations that would inform the country's forthcoming **National AI Strategy** (NAIS).

This paper analyses how Cambodia translated UNESCO's AI ethics readiness assessment (RAM) into a National AI Strategy, and what this process suggests for emerging economies designing 'ethics-by-design' AI roadmaps. Section II discusses how existing national policies and institutional arrangements laid the groundwork for the RAM exercise. Section III highlights key findings from the AI Readiness Assessment and its recommendations. Section IV then outlines the process of drafting Cambodia's National AI Strategy in 2024–2025, showing how RAM insights were translated into strategic priorities. Section V considers ways forward, including implementation of the NAIS and steps to ensure Cambodia harnesses AI for inclusive and ethical development. Throughout, the paper uses a policy-analytic lens, focusing on Cambodia's specific context – its institutions, digital policy frameworks, and international commitments. The goal is to extract Cambodia's experience that can guide its next steps and offer insights for a replicable model for emerging economies integrating ethics-by-design into AI strategy.

## II. From National Policies to Readiness Assessment

Cambodia has established a policy foundation that recognized the importance of digital transformation and, specifically, AI. **Cambodia's Digital Economy and Society Policy Framework 2021-2035**, adopted in 2021, set a 15-year vision to build a vibrant digital economy,<sup>9</sup> identifying emerging technologies like AI, data analytics, and blockchain as new drivers of growth and innovation. It called for developing digital infrastructure, nurturing digital talent, promoting research and development, and enabling e-government services – all prerequisites for effective AI deployment. Complementing this, the **Cambodia Digital Government Policy 2022-2035**, launched in 2022, outlined a roadmap to leverage digital technologies including AI and IoT to improve public service delivery and citizen quality of life.<sup>10</sup> The Digital Government Policy explicitly lists AI as a focus area for innovation in government operations and commits to developing the necessary legal frameworks, standards, and human capacity for digital governance.<sup>11</sup> In essence, these two national policies acknowledged AI as a strategic enabler for Cambodia's development, even though a dedicated AI strategy did not yet exist. They also emphasized principles like digital inclusion, trust, and security – which align with ethical AI considerations.

The **Pentagonal Strategy - Phase I** enshrined technological advancement as a pillar of Cambodia's socio-economic agenda.<sup>12</sup> This signaled to all ministries that integrating digital tools is a national priority. Indeed, in early 2025 the government delineated institutional responsibilities for AI: the Ministry of Post and Telecommunications (MPTC) was mandated to lead on AI infrastructure and governance, including coordinating AI policy and strategy development; the Ministry of Industry, Science, Technology & Innovation (MISTI) was tasked with driving AI research and development and developing an AI R&D roadmap; the Ministry

of Education, Youth and Sport (MoEYS) was assigned to lead AI-related education initiatives; and every line ministry was instructed to promote AI research, development and adoption within its own sector.<sup>13</sup> This clear coordination framework – formalized by government decision in April 2025 – ensured that by the time Cambodia moved to assess AI readiness, key institutions knew their roles.

On the international front, Cambodia's commitment to ethical AI governance strongly influenced its decision to undertake the readiness assessment. The country's endorsement of UNESCO's *Recommendation on the Ethics of AI* in 2021 bound it to uphold principles like human rights, transparency, inclusivity, and accountability in AI policy.<sup>14</sup> UNESCO subsequently developed the **Readiness Assessment Methodology (RAM)** as a tool to help countries translate those ethical principles into practice. Recognizing the opportunity, MPTC and CADT partnered with UNESCO to implement the RAM in Cambodia as a step toward a national AI action plan. This effort was also later supported by the UN Economic and Social Commission for Asia-Pacific (UNESCAP), which had launched a regional initiative to assist ASEAN countries in AI strategy development. In fact, UNESCAP convened expert consultations for the Cambodia NAIS in early 2025 and provided technical feedback on drafts.<sup>15</sup> Such multilateral support reinforced Cambodia's own policy goals, marrying local context with international best practices.

By late 2024, all the pieces were in place to officially evaluate Cambodia's AI readiness. On 12 November 2024, Cambodia launched its **AI Ethical Readiness Assessment** process (see Figure 1).<sup>16</sup> CADT led the technical work in collaboration with MPTC and guidance from UNESCO's experts. Over the following months (Nov 2024-Feb 2025), data were collected across five dimensions – legal, social, educational/scientific, economic, and technical – through extensive surveys and stakeholder

consultations.<sup>17</sup> The assessment engaged more than 300 participants from government, academia, industry and civil society, including representatives of 26 ministries and agencies.<sup>18</sup> Such a whole-of-society approach was unprecedented in Cambodia's tech policymaking; it ensured that diverse perspectives (from telecom regulators and tech startups to educators and ethicists) were captured. In February 2025, CADT and UNESCO organized validation workshops to review preliminary findings and draft recommendations with national stakeholders,<sup>19</sup> fostering consensus on priorities. The stage was then set to leverage these insights – first by finalizing the RAM country report, and next by using its findings as the evidence base for Cambodia's National AI Strategy.

Figure 1: The Implementation of the UNESCO AI RAM in Cambodia<sup>20</sup>

Research Consortium	Participated Ministries in Data Collection	Questionnaires Addressed	Consultation & Validation Workshop
19 members 5 institutions	26 ministries and agencies	176 questions 5 dimensions	300+ government, academia, industry, and civil society stakeholders

### III. The AI RAM: Findings and Recommendations

In July 2025, Cambodia completed UNESCO's Ethics of AI Readiness Assessment, becoming the fourth country in Southeast Asia to do so (after Indonesia, Thailand, and the Philippines).<sup>21</sup> The RAM report provided the first holistic portrait of Cambodia's AI ecosystem, assessing strengths and gaps in each of the five dimensions. Overall, the assessment confirmed Cambodia's strong commitment and vision for AI, while highlighting significant challenges to address before AI can be safely and widely adopted. Some key findings are summarized below by RAM dimension:

- **Legal and Regulatory**<sup>22</sup>: Cambodia has begun developing core legal instruments for the digital age, including draft laws on personal data protection and cybersecurity, alongside access-to-information and data governance policies under preparation. However, no AI-specific law or national policy was in force during the RAM period, and regulatory gaps remain in areas such as AI use in public procurement and liability for AI-related harms. Data governance is particularly weak: Cambodia scored 37 out of 100 in the 2022 Open Data Inventory (ranking 150th of 195 countries), reflecting limited data availability, fragmented government information systems, and the absence of international open data commitments. The RAM therefore recommended accelerating the adoption of key digital laws, developing AI governance guidelines, and piloting regulatory sandbox mechanisms to enable innovation while informing future AI-specific regulation.
- **Social and Cultural**<sup>23</sup>: Cambodia demonstrates strong commitment to ethical and inclusive AI, reflected in "AI for Good" initiatives (e.g., *Khmer Braille* and *Sarika*) and efforts to develop an open-source Khmer large language model to support local-language AI applications. However, inclusion challenges persist. Women remain significantly underrepresented in STEM fields (around 17 percent of graduates), and rural-urban and socio-economic digital divides continue to limit equitable access to AI benefits. Public understanding of AI remains low, as reflected in Cambodia's modest e-participation score (0.32/1) and heightened

vulnerability to online misinformation and AI-enabled abuse. The RAM therefore recommended nationwide AI literacy and awareness campaigns, integration of AI into school curricula and teacher training, and targeted measures to promote gender inclusion and context-appropriate AI tools aligned with Khmer language and local needs.

- **Scientific and Educational<sup>24</sup>:** Cambodia's education and research sector is increasingly engaging with AI, but investment in advanced skills and R&D remains limited. Digital skills development is a national priority, supported by the Digital Education Roadmap and emerging AI and data science programs in higher education, including CADT's Master's in Data Science and AI with an AI ethics component developed with UNESCO. Use of basic AI tools is growing among students and lecturers. However, structural constraints persist: national R&D spending stands at just 0.09 percent of GDP, no dedicated AI research center exists, AI research output remains modest, and many institutions lack adequate computing infrastructure and updated curricula. The RAM therefore recommended strengthening AI capacity through increased public R&D investment, establishment of a national AI research center or resource hub, expanded teacher training and curriculum integration, and enhanced international research collaboration to accelerate skills development and innovation.
- **Economic<sup>25</sup>:** Cambodia views AI as a key driver of productivity and economic diversification, but evidence of adoption and impact remains limited. AI applications are emerging in government pilots (such as Khmer-language translation and predictive analytics) and in private sectors including finance, agriculture, tourism, and e-commerce. However, AI integration in the economy is still at a nascent stage. Systematic data on AI adoption by firms, AI-related employment, and private sector R&D spending are largely absent, and high-tech exports account for only 6.6 percent of manufactured exports. To accelerate AI-enabled growth, the RAM recommended sector-specific AI action plans, targeted incentives for business and SME adoption, improved data collection on the digital economy, and the use of innovation sandboxes and start-up incubation to strengthen the domestic AI ecosystem.
- **Technical and Infrastructural<sup>26</sup>:** Cambodia has achieved strong basic digital connectivity—mobile broadband coverage exceeds 95 percent, internet penetration is above 70 percent, and electricity access is nearly universal—and has introduced platforms such as CamDX to support digital government integration. However, AI-enabling infrastructure remains limited. High-performance computing, large-scale data storage, and domestic cloud capacity are scarce; the country has only two Tier-III data centers, with a National Data Center still under development. Constraints in computing power, limited Khmer-language data for AI applications, fragmented government information systems, and low participation in international AI standards bodies further hinder AI development. To address these gaps, the RAM recommended investing in shared AI infrastructure (including cloud and high-performance computing), strengthening data governance through an Open Data Policy and data-sharing framework, establishing a National AI Lab as an experimentation sandbox, and increasing engagement in international AI standard-setting initiatives.

Beyond diagnosing gaps, the RAM issued more than 30 actionable recommendations that effectively shaped Cambodia's National AI Strategy. Core priorities included establishing a multi-stakeholder AI governance mechanism; fast-tracking adoption of the National AI Strategy with clear objectives and evaluation mechanisms; accelerating enabling legislation, particularly the Personal Data Protection Law and access-to-information frameworks; and developing national and sector-specific ethical AI guidelines grounded in transparency, fairness, and human oversight.<sup>27</sup> The RAM also emphasized investment in AI education and literacy, strengthening digital infrastructure and data ecosystems (including open data and Khmer-language resources), and boosting research, innovation, and international collaboration through dedicated programs and sandbox environments.<sup>28</sup> As UNESCO underscored, the RAM provides a "clear roadmap" for embedding ethical principles into Cambodia's AI development from the outset.<sup>29</sup> The Cambodian government demonstrated early ownership of these recommendations by translating them directly into the NAIS drafting process described below.

#### IV. The Drafting of Cambodia's National AI Strategy

Even as the RAM was being finalized in the first half of 2025, Cambodia had already initiated the drafting of its National AI Strategy (NAIS) in parallel. The strategy development was a year-long, iterative process led by MPTC's technical working group, with milestones carefully synchronized with the RAM findings and broader consultation efforts. This section outlines how the NAIS was formulated, including key inputs, stakeholder engagement, and the resulting strategic framework.

Cambodia's National AI Strategy development formally began in early 2024. On 5 March 2024, the Royal Government approved MPTC's proposal to establish a technical working

group to monitor AI development, followed by the establishment of an Artificial Intelligence and Data Science Research and Development Working Group in April 2024.<sup>30</sup> Between April 2024 and February 2025, the Working Group conducted numerous internal meetings and produced an initial draft informed by preliminary RAM findings, international AI strategies, and global frameworks from UNESCO and the OECD.<sup>31</sup> Regional guidance, including the ASEAN Guide on AI Governance and Ethics adopted in early 2024, was also integrated to ensure alignment with ASEAN priorities. Overall, the initial draft sought to adapt international best practices to Cambodia's national development and digital policy context.

**Expert consultations:** Following completion of the initial draft, Cambodia undertook external expert consultations to refine the strategy. From 19 February to 18 March 2025, UNESCAP facilitated virtual consultations with over 30 AI experts from across the Asia-Pacific region, whose feedback informed a revised draft completed in late March 2025.<sup>32</sup> This was followed by a two-day Deep Dive Workshop in Phnom Penh on 7–8 April 2025, co-organized by MPTC, UNESCO, and UNESCAP, which enabled a detailed review of the draft and direct integration of validated RAM findings, particularly on ethical AI, inclusion, and prioritization of strategic actions. The process resulted in draft of NAIS reflecting a more focused and evidence-based strategy.

**Government and public consultations:** In May 2025, MPTC conducted capacity-building and inter-ministerial consultations to integrate sectoral priorities into the draft strategy. These inputs culminated in NAIS Draft v4 by 20 May 2025, which was released for public consultation on 11 June 2025.<sup>33</sup> Although the online consultation window was narrow, it is one of the first open online public consultations for a national policy document, demonstrating the ministry's commitment to a multi-stakeholder approach in digital policy

development. A national consultation workshop held on 1–2 July 2025, alongside the launch of the UNESCO RAM report, brought together over 450 stakeholders from government, the private sector, civil society, and development partners to review the draft.<sup>34</sup> Feedback from this process was incorporated into NAIS Draft v5, which served as the near-final version of the strategy.

**Strategic framework of NAIS:** it is a comprehensive document charting Cambodia's AI development path for the coming years to 2030. It opens with an ambitious vision: *"Promote artificial intelligence adoption and development, for increasing productivity and generating added value, to improve citizens' quality of life and enhance sustainable and inclusive socio-economic development."*<sup>35</sup> This vision captures the dual aim of the NAIS – to harness AI for economic and governance gains, and to do so in a way that benefits all Cambodians. From this, the strategy derives five overarching goals: (1) upgrade citizens' digital capabilities; (2) harness AI for digital government; (3) promote economic development through AI; (4) strengthen the national AI ecosystem (infrastructure and innovation); and (5) manage the risks of AI to ensure it is safe and ethical.

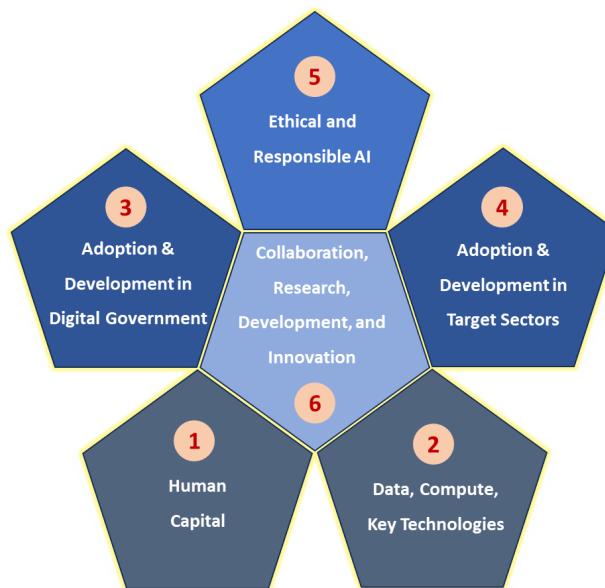


Figure 2: Six Strategic Priority of the Draft NAIS<sup>36</sup>

To realize these goals, the NAIS is built around **six Strategic Priorities** (as illustrated in Figure 2), each with a dedicated set of strategies and initiatives:

- **Strategic Priority 1: Human Resource Development**<sup>37</sup> – focusing on cultivating AI talent and literacy. This includes a national AI literacy campaign to raise general awareness, integration of AI topics into general and technical education, “train-the-trainer” programs, and scholarships. A flagship objective is to produce “1,000 AI talents in 5 years,” through new degree and certification programs in AI and data science, as well as encouraging STEM education in schools. Leadership training (for government and business leaders to understand AI) is also planned. This pillar directly addresses the skills gap identified in the RAM.
- **Strategic Priority 2: Data, Infrastructure, and Key Technologies**<sup>38</sup> – aiming to bolster the foundational digital infrastructure and resources needed for AI. Key actions include

promoting cloud adoption in both public and private sectors to increase affordable access to computing power; developing and implementing a national Data Governance Policy and accelerating the Open Data initiative to ensure data availability and interoperability; creating large Khmer language datasets and knowledge graphs to fuel local AI applications; and investing in key technologies such as a Khmer Natural Language Processing and Speech technologies, and even exploring the development of a Khmer Large Language Model (LLM). A major proposal is to establish a National AI & Data Science Research Center, which would serve as a hub for R&D and provide shared high-performance computing resources.

- **Strategic Priority 3: Promoting AI Adoption and Development for Digital Government<sup>39</sup>** – aligns with the Digital Government Policy and proposes the use of AI to improve public services and government operations. It focuses on identifying and piloting priority government use cases—such as AI-enabled public service chatbots, data analysis for policymaking, and administrative automation—and on strengthening AI capacity within the civil service, including the Digital Government Technical Team at MPTC and technical units in line ministries. By scaling AI applications in e-government, including existing pilots such as automated translation on government websites, the pillar aims to enhance efficiency and build public trust in AI. Oversight is provided by the Digital Government Committee under MPTC to ensure coherence with digital government reforms.
- **Strategic Priority 4: Promoting Sectoral Adoption and Development<sup>40</sup>**

– recognizes that AI's impact extends beyond the technology sector and targets eight priority areas: agriculture, health, education, finance, manufacturing/industry, tourism, transport and logistics, and Micro, Small, and Medium Enterprises (MSMEs) including e-commerce. For each sector, the NAIS identifies potential AI applications and sector-specific strategies, such as smart farming and crop analytics in agriculture; medical imaging, telemedicine, and public health analytics in health; AI-enhanced learning tools and education management systems; chatbots, fraud detection, and inclusive fintech in finance; automation, predictive maintenance, and quality control in manufacturing; AI-enabled tourism services and planning analytics; and route optimization and traffic management in transport. Line ministries and sector regulators are expected to integrate AI into their respective sectoral plans, with public-private partnerships encouraged to pilot applications and enable data sharing, ensuring broad-based AI-driven development across the economy (see Figure 3).

- **Strategic Priority 5: Promoting Ethical and Responsible AI<sup>41</sup>** – focuses on managing AI risks and ensuring alignment with societal values, in line with the RAM and UNESCO's ethics framework. It envisages the development of a National AI Governance Framework and National Guidelines on AI Ethics and Governance, drawing on ASEAN and UNESCO principles to guide government and industry on fairness, transparency, privacy, and accountability. The pillar reinforces the urgency of enacting and enforcing the Personal Data Protection Law and proposes technical compliance tools such as algorithmic audits,

AI impact assessments, and mechanisms for identifying AI-generated content. It also emphasizes sector-specific risk management, the establishment of an oversight body or multi-stakeholder AI advisory council, and public awareness campaigns on AI risks and benefits to ensure responsible adoption as AI use expands.

- **Strategic Priority 6: Promoting Collaboration, Research, Development and Innovation<sup>42</sup>** – highlights cross-cutting enablers to sustain Cambodia's AI ecosystem through whole-of-government coordination, public-private-academic collaboration, and international engagement, including within ASEAN and global AI initiatives. It sets out actions to increase AI-related R&D funding, encourage private-sector innovation and startups, and establish research facilities such as an AI research center and an AI Resource Center or knowledge hub. The pillar also promotes international research partnerships and exchange programs, and assigns responsibility for monitoring and evaluating NAIS implementation, including the formation of an inter-ministerial implementation committee and annual progress reviews.

*Figure 3: Illustrative AI Applications in Priority Sectors  
(Strategic Priority 4 – Promoting Sectoral Adoption and Development)<sup>43</sup>*



The NAIS document details 42 strategic measures across these six strategic priorities. Each measure corresponds to specific programs or policy actions, complete with designated lead agencies, timelines, and expected outputs. This level of detail is meant to facilitate implementation. The NAIS also outlines an implementation and governance mechanism to coordinate these efforts. It proposes that the Digital Government Committee coordinates its implementation and M&E, supported by a subordinate interministerial working group established particularly for AI. Relevant units of MPTC would provide technical and administrative support for these higher bodies. Furthermore, the strategy leverages the existing **AI focal point** nominated by all members of the DGC at the Chair's request, who will liaise with the Committee and drive sector-specific AI initiatives. Furthermore, the strategy calls for each ministry to designate an **AI focal point** (a technical expert or senior official) who will liaise with the Committee and drive sector-specific AI initiatives.

**Finalization and current status:** Minor revisions were made, resulting in NAIS Version 6 and presented in the 8<sup>th</sup> Digital Government Committee Meeting.<sup>44</sup> It is expected to serve as a roadmap through 2030. The journey of developing this strategy – from initial readiness assessment to a full-fledged policy document – has been characterized by a high degree of stakeholder engagement, evidence-based planning, and alignment with both national goals and international principles. Cambodia's NAIS is thus poised to be one of the first in the Least Developed Countries (LDC) group and a model for how to proactively and responsibly embrace AI. Dr. Phinith Chanthalangsy of UNESCO's regional office lauded Cambodia's completion of the RAM and progress on the NAIS as "a major milestone in its digital transformation journey," noting it offers a roadmap for ethical AI at a critical time as Cambodia aims for LDC graduation and deeper digital economy integration.<sup>45</sup> The focus now shifts to turning this strategy from paper to practice.

## V. Ways Forward

Cambodia's National AI Strategy (NAIS) provides a clear strategic direction; however, its success relies on the immediate prioritization of actions that translate this framework into coordinated practice.

**First, finalize adoption and establish an implementation mechanism.** Because the NAIS is not yet formally adopted, the immediate priority is to complete the endorsement process and operationalize the coordination structure envisioned in the strategy. This includes empowering MPTC to convene an inter-ministerial implementation arrangement, clarify roles across ministries, and ensure political follow-up so the strategy does not stall at the drafting stage.

**Second, operationalize Ethical and Responsible AI governance.** As emphasized by the RAM and UNESCO's ethics framework, Cambodia should move early on the NAIS's responsible AI pillar by developing practical governance instruments—such as national AI ethics and governance guidelines, tools for risk and compliance (e.g., impact assessments or audit mechanisms), and sector-specific risk management approaches. Advancing core enabling legislation and enforcement capacity for data protection remains a central part to this priority.

**Third, drive AI adoption in priority sectors through targeted pilots and partnerships.** To convert strategy into visible outcomes, Cambodia should focus on adoption in key sectors identified in the NAIS (including digital government and priority economic and social sectors) through well-sscoped pilots and public-private collaboration. Early "quick wins" can build institutional learning and public trust, while generating evidence for scaling, budgeting, and future regulatory refinement. It is extremely advisable engage the private sector and target groups in the population, especially people at the grassroots level of socio-economic activities, to participate in

AI adoption exercise as they have the richest understanding of the problems, for which AI may form part of the solution. AI is a means, not an end. A pragmatic, iterative, and collaborative process is needed to avoid inadvertently turning AI into a solution in search of a problem and to unlock its full potential to create value.

## References

1. Stanford Institute for Human-Centered Artificial Intelligence (HAI), *Artificial Intelligence Index Report 2025* (Stanford, CA: Stanford University, 2025), 4, <https://hai.stanford.edu/ai-index/2025-ai-index-report>.
2. Organisation for Economic Co-operation and Development (OECD), "OECD AI Policy Observatory: National AI Strategies," OECD. ai, accessed December 20, 2025, <https://oecd.ai/en/dashboards/policy-initiatives>.
3. Royal Government of Cambodia (RGC), *Pentagonal Strategy – Phase I: Growth, Employment, Equity, Efficiency, and Sustainability* (Phnom Penh: RGC, 2023), 8
4. Royal Government of Cambodia (RGC), *Digital Economy and Society Policy Framework 2021–2035* (Phnom Penh: RGC, 2021), <https://www.mptc.gov.kh/sites/default/files/2021-06/Digital%20Economy%20and%20Society%20Policy%20Framework%202021-2035.pdf>.
5. Royal Government of Cambodia (RGC), *Cambodia Digital Government Policy 2022–2035* (Phnom Penh: RGC, 2022), <https://www.mptc.gov.kh/sites/default/files/2022-06/Cambodia%20Digital%20Government%20Policy%202022-2035.pdf>.
6. UNESCO, "UNESCO Adopts First Global Standard on the Ethics of Artificial Intelligence," UNESCO, November 25, 2021, <https://www.unesco.org/en/articles/unesco-adopts-first-global-standard-ethics-artificial-intelligence>.
7. UNESCO, "Cambodia Launches AI Readiness Assessment Report to Guide Ethical and Inclusive Digital Transformation," UNESCO, July 9, 2025, <https://www.unesco.org/en/articles/cambodia-launches-ai-readiness-assessment-report-guide-ethical-and-inclusive-digital-transformation>.
8. UNESCO, *Cambodia: Summary AI Readiness Assessment Report* (Phnom Penh: UNESCO, 2025), [Page Number], <https://articles.unesco.org/sites/default/files/medias/fichiers/2025/07/CAMBODIA%20-%20Summary%20AI%20Readiness%20Assessment%20Report%20-%20ENG%20.pdf>.
9. RGC, *Digital Economy and Society Policy Framework 2022-2035*, v.
10. RGC, *Cambodia Digital Government Policy 2022–2035*, 14.
11. *Ibid*, 36-49
12. RGC, *Pentagonal Strategy – Phase I*, 24.
13. UNESCO, *Cambodia: Artificial Intelligence Readiness Assessment Report* (Paris: UNESCO, 2025), 15, <https://unesdoc.unesco.org/ark:/48223/pf0000394553>.
14. UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (Paris: UNESCO, 2021), <https://unesdoc.unesco.org/ark:/48223/pf0000380455>.
15. Ministry of Post and Telecommunications (MPTC), "Press Release on Process of Artificial Intelligence Governance," [ai.gov.kh](https://ai.gov.kh/press-release-on-process-of-artificial-intelligence-governance/), April 28, 2025, <https://ai.gov.kh/press-release-on-process-of-artificial-intelligence-governance/>.
16. *Ibid*.
17. UNESCO, "Cambodia Launches AI Readiness Assessment Report."
18. *Ibid*
19. MPTC, "Press Release on Process of AI Governance."
20. UNESCO, *Cambodia: Summary AI Readiness*, 2.
21. UNESCO, "Cambodia Launches AI Readiness Assessment Report."
22. UNESCO, *Cambodia: AI Readiness Assessment Report*, 17-23.

**Mr. Chanty Pisal**

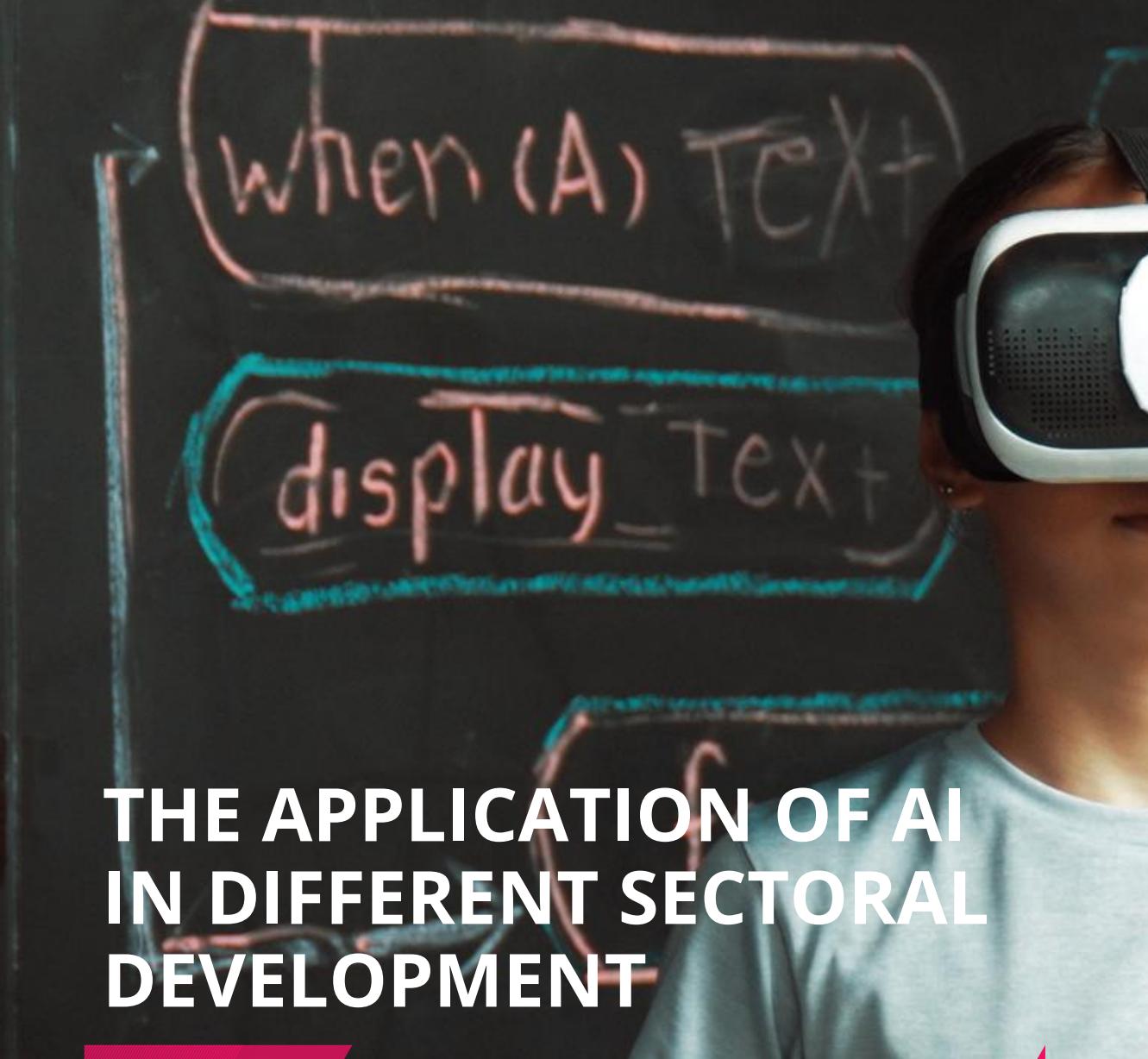
*Mr. Chanty Pisal is the Director of the Digital and Innovation Policy Research Center at the Cambodia Academy of Digital Technology (CADT). He is also a member of the UNESCO AI Ethics Experts Without Borders network and contributed significantly to the implementation of Cambodia's UNESCO AI Readiness Assessment Methodology (RAM) report. He also serves on the drafting team for Cambodia's Digital Skill Development Roadmap and a project coordinator of the UNESCO Internet Universality Indicators assessment for Internet development in Cambodia.*

*His extensive professional background includes serving as Vice Director of the ASEAN ICT Center and as an International Cooperation Officer at the Ministry of Post and Telecommunications, where he managed portfolios for the WTO, ASEAN, and Mekong Cooperation. Mr. Pisal holds a Master's in Public Policy from the Australian National University and a Master's in International Law from the University of Paris 8. His research focuses on the intersection of the digital economy, AI law and policy, cybersecurity, and ASEAN regional integration.*

**Mr. Taing Bunhong**

*Mr. Taing Bunhong is the Director of Policy and Legal Affairs at the General Secretariat of the Digital Government Committee, and Vice-Chair of AI and Data Science R&D Working Group of the Ministry of Post and Telecommunications (MPTC). Before his current position, Bunhong worked as the officer at the Office of the Council of Ministers (OCM).*

*Bunhong earned his Bachelor's degree in law from the American University of Phnom Penh (AUPP) in 2017 and further pursued an LLM in Drafting Legislation, Regulation and Policy at School of Advanced Study, University of London in 2020 as a Chevening scholar.*



# THE APPLICATION OF AI IN DIFFERENT SECTORAL DEVELOPMENT

**52**

AI in Higher Education: Attitudes and Acceptance Among University Students in Cambodia

*Borey Be*

**78**

AI Meets Public Data: A Prompt-Based Agent for Intuitive Analysis of Government Datasets

*Rina Buoy, Nguonly Taing, and Marry Kong*

**88**

Training the Trainers: Scaling AI Annotation Through Inclusive Workforce Development

*Songheng Tieng, Sokheng Pok, and Chantola Khet*



# AI in Higher Education: Attitudes and Acceptance Among University Students in Cambodia



## I. Introduction

### 1. Understanding Artificial Intelligence Integration in Higher Education

Artificial Intelligence (AI) is rapidly becoming a foundation of technological innovation across various industries, including education. In higher education, AI has the potential to transform teaching and learning by providing personalized educational experiences, automating administrative processes, and utilizing advanced data analytics to monitor and enhance the student performance.<sup>1</sup> However, the successful integration of AI technologies in educational systems is not merely dependent on the availability of technology but also on the attitudes, perceptions, and acceptance of these technologies by key stakeholders such as students and teachers.<sup>2</sup>

Understanding why users accept or reject technologies is essential to the effective implementation of AI. The Technology Acceptance Model (TAM), developed by Davis (1989), has been generally investigated to study how users come to accept and use the technology.<sup>3</sup> The model posits that perceived usefulness (PU) and perceived ease of use (PEOU) are primary determinants of technology adoption. PU refers to the degree to which a person believes that using a particular system would enhance their job performance, while PEOU refers to the extent to which a person believes that using the system would be free of effort.<sup>4</sup> Building on TAM, the Unified Theory of Acceptance and Use of Technology (UTAUT) proposed by

Venkatesh et al. (2003) incorporates additional constructs such as social influence (SI) and facilitating conditions (FC).<sup>5</sup> These models have been extensively validated across various contexts and technologies, including in educational settings.<sup>6</sup> In the context of AI in education, emerging concerns about AI's ethical implications, particularly related to data privacy, fairness, and algorithmic bias, are increasingly recognized as critical factors shaping user acceptance. Ethical concerns can significantly influence whether students and teachers are willing to embrace AI tools.<sup>7</sup> Even though extensive research exists on AI adoption in developed countries, there is a notable lack of empirical studies focusing on developing regions such as Southeast Asia and specifically Cambodia. Addressing this gap is crucial for understanding how AI can be effectively and responsibly integrated in diverse educational contexts.

### 2. National AI Strategy and Challenges for Ethical AI Integration in Cambodia's Higher Education

Cambodia is actively developing a national framework for AI that aims to balance innovation with ethical and responsible use. The Royal Government, through the Ministry of Post and Telecommunication (MPTC), is finalizing its first National Artificial Intelligence Strategy. This draft strategy, shaped by stakeholder consultations, technical workshops, and international collaboration, focuses on six strategic priorities: human resource development; data, infrastructure,

and technologies; AI for digital government; sectoral adoption; ethical and responsible AI; and research and innovation.<sup>8</sup> With 41 strategic measures, the government seeks to position Cambodia as a skilled AI adopter in the short term and an active participant in global AI development in the longer term.<sup>9</sup>

Complementing this strategy is the Cambodia Digital Government Policy 2022–2035, which guides the country's broader digital transformation efforts. This policy seeks to improve public service delivery and citizen trust through coordinated digitalization, with AI piloted in key sectors including education, health, agriculture, banking, and public administration.<sup>10</sup> Cambodia has also engaged in UNESCO's Ethics of AI Readiness Assessment Methodology (RAM) and has officially launched its completed country report, a major milestone in its digital transformation journey. This helps to ensure that upcoming policies embed ethical, rights-based, and inclusive digital transformation principles. As of mid-2025, Cambodia is in an active policy-making stage, AI systems are not yet comprehensively regulated, but primary legal instruments and guidelines are in development to address concerns over data protection, transparency, and responsible AI innovation.<sup>11</sup>

Despite these policies, there are remaining challenges in integrating AI effectively within Cambodia's higher education sector. The key challenges in AI usage and practices in Cambodia's higher education include inadequate technological infrastructure, limited access to digital resources, and varying levels of digital literacy among students, and gaps in AI literacy and training.<sup>12</sup> Moreover, the fast growth of AI development raises ethical and privacy concerns that need careful consideration, particularly in contexts with underdeveloped regulatory frameworks.<sup>13</sup> It is costly for many students to access the internet, which limits the use of AI-driven learning tools effectively. There is also a reported shortage of localized AI tools and resources in the

Khmer language,<sup>14</sup> which hinders the practical usability and inclusiveness. This digital divide risks worsening existing inequalities within the education system and limits the full potential of AI as a tool for learning enhancement.<sup>15</sup>

Finally, the policy and strategic frameworks for AI in education remain in progress. The need for a socially responsive, equitable digital strategy that considers Cambodia's socio-economic context is important to promote responsible AI adoption and data governance in the education sector. Stakeholder collaboration among government, educational institutions, and international partners is emphasized as key to overcoming these challenges, balancing the approach between technology and human elements in higher education.<sup>16</sup>

### **3. Purposes of the Chapter**

The purpose of this chapter is to investigate and understand the various factors that shape students' attitudes and acceptance of AI technologies within Cambodia's higher education context. Specifically, the chapter seeks to answer the following research questions:

- What are the perceptions of students regarding the usefulness and ease of use of AI technologies in higher education?
- How do social influence, facilitating conditions, and ethical concerns impact the behavioral intention to use AI technologies?
- To what extent do students in Cambodian higher education institutions use AI technologies in their academic activities?

The chapter uncovers how students' perceptions of AI's usefulness and ease of use influence their willingness to adopt and engage with AI tools in learning and academic pursuits. Additionally, it seeks to examine the roles of social influence, facilitating conditions,

and ethical concerns in shaping students' behavioral intentions toward AI use. Finally, the study intends to measure the actual extent of AI technology usage by students in Cambodian higher education institutions.

## II. Methods

### 1. Research Design

This chapter employs a quantitative research design to examine the attitudes and acceptance of AI technologies among students in higher education institutions in Cambodia. A quantitative approach is particularly suited for this chapter as it allows for the systematic collection and analysis of data from a large sample, providing a more generalizable understanding of the factors influencing AI acceptance.<sup>17</sup> The chapter is cross-sectional, gathering data at a single point in time to explore the relationships between key variables.

### 2. Participants

The chapter sample consists of 256 students (149 females) from various higher education institutions (HEIs) in Cambodia. The selection of participants was based on a random sampling technique, ensuring representation across different universities and academic disciplines. This approach was chosen to capture a diverse range of perspectives and experiences with AI, which is crucial for understanding the general attitudes and acceptance of these technologies in the Cambodian higher education context.

### 3. Data Collection

Data was collected through a structured survey questionnaire designed to measure the key dimensions of TAM. The survey was distributed electronically via emails, telegram groups and in the classrooms facilitated by the teachers.

The survey instrument was developed based on the existing literature. For instance, PU and PEOU were measured using items adapted from Davis (1989),<sup>18</sup> while SI and FC were measured using items adapted from Venkatesh et al. (2003).<sup>19</sup> Participants were asked to respond to each item using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The reliability analyses show high internal consistency for all constructs (Cronbach's alpha above 0.90), confirming that items within each scale consistently measure the same underlying concept, thus providing confidence in the validity of these results.

*Table 1: Reliability Score*

Scale	Mean	SD	Cronbach's $\alpha$
PU	3.473	0.952	0.917
PEOU	3.402	0.924	0.912
ATT	3.449	0.967	0.922
BI	3.407	0.976	0.938
SI	3.304	0.926	0.925
FC	3.19	0.96	0.924
EPC	3.418	0.949	0.917

*Note: PU = Perceived Usefulness, PEOU = Perceived Ease of Use, ATT = Attitude toward AI, BI = Behavioral Intention, SI = Social Influence, FC = Facilitating Conditions, EPC = Ethical and Privacy Concern*

### 4. Data Analysis

Data analysis was conducted using Jamovi version 2.5.5. Descriptive statistics were first computed to provide an overview of students' overall attitudes and perceptions toward AI. The Shapiro-Wilk test was used to assess normality, and although results indicated significant deviations ( $p < .001$ ), the large sample size ( $N = 256$ ) allowed for the assumption of approximate normality in accordance with the Central Limit Theorem.<sup>20</sup> The independent samples t-tests were performed to compare

male and female students' mean scores on key constructs. In addition, one-way analysis of variance (ANOVA) was conducted to test significant differences among academic disciplines. Where significant effects were found, Tukey HSD post hoc tests were applied to identify specific group differences.<sup>21</sup>

Boxplots and histograms were generated to visually assess the distribution of responses and to measure the central tendency and outliers for each construct. These visualizations complemented the numerical analysis by illustrating how responses were clustered around the upper end of the Likert scale. To further explore response patterns and categorize students based on their levels of AI acceptance, a K-means cluster analysis was conducted using key variables (PU, PEOU, ATT, BI, SI, FC, EPC, and AU). ANOVA tests were subsequently used to confirm significant mean differences between clusters across all constructs. Finally, hierarchical multiple regression analysis was employed to examine the predictive relationships among the constructs. The model assessed how PU, PEOU, ATT, SI, FC, and EPC influenced students' behavioral intention to use AI, and how behavioral intention influenced. The findings from these analyses collectively identified the key factors shaping students' acceptance and utilization of AI technologies.



### III. Findings

#### 1. Students' Perceptions and Attitudes Toward AI in HEIs

The descriptive statistics indicate that students generally perceive AI technologies as beneficial and moderately easy to use. For PU, students agreed that AI enhances academic performance ( $M = 3.48$ ,  $SD = 1.02$ ), improves educational quality ( $M = 3.47$ ,  $SD = 1.07$ ), supports personalized learning ( $M = 3.50$ ,  $SD = 1.08$ ), and aids decision-making in academic activities ( $M = 3.45$ ,  $SD = 1.08$ ). Similarly, PEOU ratings show moderate agreement that AI is user-friendly ( $M = 3.46$ ,  $SD = 0.98$ ) and relatively easy to integrate into learning purposes ( $M = 3.41$ ,  $SD = 1.05$ ), though respondents felt slightly less certain about the mental effort required ( $M = 3.32$ ,  $SD = 1.04$ ). The results reflect an overall positive view of AI's usefulness and accessibility, which is essential for adoption.

Students expressed a generally favorable attitude toward AI, highlighting excitement for its potential to enhance education ( $M = 3.52$ ,  $SD = 1.05$ ) and belief in its important role in modern learning environments ( $M = 3.51$ ,  $SD = 1.16$ ). The Behavioral Intention to use AI was also moderately positive, with intentions to explore more AI tools ( $M = 3.42$ ,  $SD = 1.09$ ), recommend AI use to others ( $M = 3.44$ ,  $SD = 1.07$ ), and continue using AI despite challenges ( $M = 3.39$ ,  $SD = 1.03$ ). Regarding SI, students moderately agreed that peers, colleagues, and leadership support AI adoption ( $M = 3.30$ ,  $SD = 0.926$ ), though the endorsement is not overwhelming. EPC was notable, with students agreeing that AI should be used responsibly ( $M = 3.52$ ,  $SD = 1.08$ ) and transparently, free from bias ( $M = 3.37$ ,  $SD = 1.09$ ).

Table 2: Descriptive Statistics (N=256)

Items	Mean	SD	Percentiles		
			25th	50th	75th
Using AI technologies enhances my academic performance/teaching effectiveness. (PU1)	3.48	1.02	3	4	4
AI technologies improve the quality of education in my institution. (PU2)	3.47	1.07	3	4	4
The use of AI in education allows for more personalized learning/teaching experiences. (PU3)	3.50	1.08	3	4	4
AI technologies help in making more informed decisions in academic/teaching activities. (PU4)	3.45	1.08	3	4	4
Learning to use AI technologies is easy for me. (PEOU1)	3.42	1.09	3	3	4
AI technologies are user-friendly and easy to interact with. (PEOU2)	3.46	0.98	3	4	4
I find it easy to integrate AI technologies into my academic/teaching practices. (PEOU3)	3.41	1.05	3	3	4
Using AI technologies does not require a lot of mental effort. (PEOU4)	3.32	1.04	3	3	4
I have a positive attitude toward the use of AI technologies in education. (ATT1)	3.41	1.06	3	3	4
I believe that AI can play an important role in modern education. (ATT2)	3.51	1.16	3	4	4
I am excited about the potential of AI to enhance the learning/teaching experience. (ATT3)	3.52	1.05	3	4	4
I believe that the advantages of using AI in education outweigh any potential disadvantages. (ATT4)	3.36	1.02	3	3	4
I intend to use AI technologies in my future academic/teaching activities. (BI1)	3.38	1.07	3	3	4
I plan to explore more AI tools and resources to enhance my academic/teaching practices. (BI2)	3.42	1.09	3	4	4
I am likely to recommend the use of AI technologies to others in my institution. (BI3)	3.44	1.07	3	4	4
I will continue to use AI technologies even if I encounter challenges. (BI4)	3.39	1.03	3	3	4
I believe that AI technologies will help me achieve better educational outcomes. (PE1)	3.46	1.07	3	4	4
AI can help improve efficiency in my academic/teaching activities. (PE2)	3.36	1.09	3	3	4

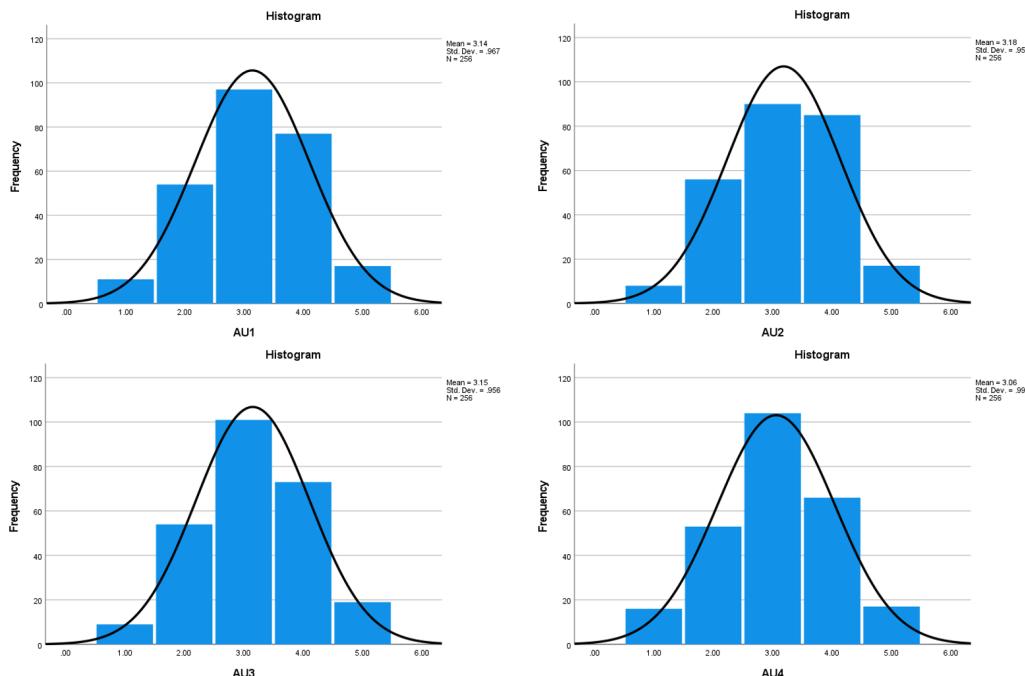
Using AI technologies will allow me to accomplish tasks more quickly. (PE3)	3.55	1.08	3	4	4
AI technologies can enhance the learning/teaching experience by providing new insights. (PE4)	3.43	1.02	3	4	4
It is easy for me to become proficient in using AI technologies. (EE1)	3.33	1.05	3	3	4
I find AI technologies easy to use. (EE2)	3.41	1.05	3	3	4
Using AI technologies does not require extensive technical knowledge. (EE3)	3.21	1.06	2	3	4
I feel confident in my ability to use AI technologies effectively. (EE4)	3.25	1.08	2.75	3	4
People important to me think I should use AI technologies in my academic/teaching activities. (SI1)	3.31	1.01	3	3	4
My peers and colleagues encourage the use of AI technologies in education. (SI2)	3.31	1.02	3	3	4
The leadership at my institution supports the integration of AI technologies in education. (SI3)	3.30	1.04	3	3	4
There is a positive attitude toward AI technologies within my institution. (SI4)	3.30	1.02	3	3	4
My institution provides adequate training on the use of AI technologies. (FC1)	3.17	1.09	2	3	4
I have access to the necessary resources to use AI technologies in my academic activities. (FC2)	3.25	1.05	3	3	4
The technological infrastructure at my institution supports the use of AI technologies. (FC3)	3.15	1.04	3	3	4
I can get help from technical support when I encounter problems with AI technologies. (FC4)	3.20	1.07	2.75	3	4
I am concerned about the privacy of my data when using AI technologies. (EPC1)	3.41	1.05	3	3.5	4
I believe that AI technologies should be transparent and free from bias. (EPC2)	3.37	1.09	3	4	4
The ethical implications of using AI in education concern me. (EPC3)	3.37	1.02	3	3	4
I believe that AI technologies should be used responsibly in education. (EPC4)	3.52	1.08	3	4	4

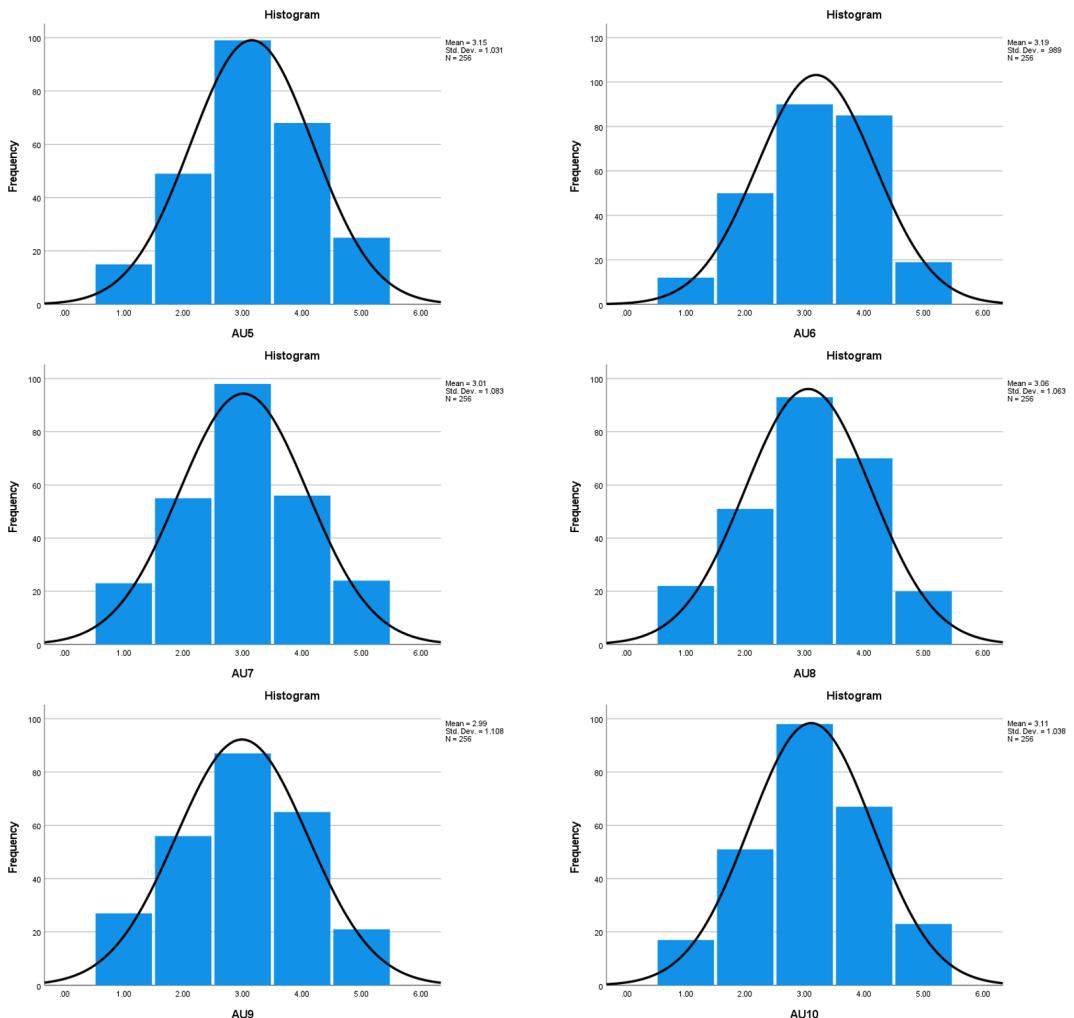
Note: Scale (1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree)  
 Minimum = 1; Maximum = 5

The 50th (median) and 75th percentiles were identical for several items (both = 4), indicating that at least half of the students selected “Agree.” This clustering of responses around the upper category reflects a generally favorable attitude toward AI use, accompanied by limited variability among high-end responses. The overlap between the median and upper quartile also suggests a mild ceiling effect, whereby students tended to agree with positive statements about AI but seldom chose the highest response option (“Strongly Agree”). These results imply that students perceived AI as beneficial and relevant, yet their responses remained moderate rather than strongly affirmative.

## 2. AI Usage in Cambodian Higher Education

The analysis of AI-based tool usage among Cambodian university students indicates moderate engagement. On a 5-point scale (1 = Never; 5 = Very Often), mean scores for ten usage items range from 2.99 to 3.19, reflecting usage from “occasionally” to “often,” but not “very often.” Activities such as general educational purposes and AI-related topic exploration show slightly higher engagement (means > 3.15), while advanced or specialized tasks remain less frequent. Most students report occasional to frequent use (25–40%), with very frequent use (<10%) limited and rare/never use (3–10%) minimal. These findings suggest growing acceptance of AI tools in Cambodian higher education, though adoption is largely non-routine. Barriers such as access, confidence, and institutional policies constrain widespread use. To foster consistent integration, investments in infrastructure, training, and awareness of diverse AI applications are essential.





### 3. Group Differences in Attitudes and Acceptance of AI

#### 3.1. Gender Comparison

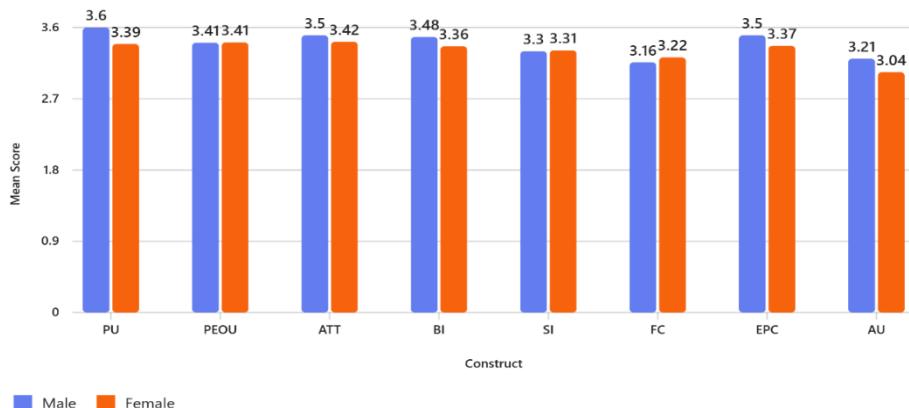
The independent samples t-test was conducted to compare male and female students' attitudes and acceptance of AI in higher education across eight constructs: PU, PEOU, ATT, BI, SI, FC, EPC, and AU. Results indicate no statistically significant differences between genders ( $p > 0.05$  for all constructs). Although male students reported slightly higher mean scores for PU, ATT, BI, EPC, and AU, and the female students scored marginally higher on PEOU, SI, and FC, these differences were not significant.

Table 3: Independent Samples t-Test Comparing Male and Female Students

Construct	Levene's Test F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
PU	0.003	0.957	1.733	252	0.084	0.20906	0.12066
PEOU	3.565	0.06	-0.033	252	0.973	-0.00393	0.1174
ATT	0.621	0.432	0.654	252	0.514	0.08054	0.12322
BI	0.433	0.511	0.951	252	0.343	0.11854	0.12467
SI	3.793	0.053	-0.088	252	0.93	-0.0104	0.11848
FC	3.466	0.064	-0.505	252	0.614	-0.06195	0.12261
EPC	0.119	0.731	1.084	252	0.28	0.13087	0.12079
AU	0.811	0.369	1.606	252	0.11	0.17014	0.10596

This suggests that gender does not play a meaningful role in shaping attitudes or acceptance of AI among Cambodian university students. Therefore, AI adoption strategies should focus on overall usability and perceived benefits rather than gender-specific approaches.

Figure 1: Bar chart comparing mean scores by gender for each construct



Across eight constructs in figure 1, male and female students show broadly similar perceptions with small gaps (mostly under 0.2 points). Both groups rate PU and EPC highest, indicating strong recognition of usefulness. Female students score PEOU, ATT, SI, and (FC slightly higher than males, suggesting usability and support factors resonate marginally more with female students. Conversely, male students score higher on BI and AU, implying a modestly stronger tendency to use the system in practice.

### 3.2. Discipline Comparison of AI Attitude and Acceptance

A one-way ANOVA was conducted to examine differences in attitudes and acceptance of AI in higher education across four disciplines: Social Science, Arts and Humanities, Natural Science, and Formal Science. Results revealed no statistically significant differences among disciplines for PU, PEOU, ATT, BI, SI, FC, and EPC, with all p-values exceeding .05. However, actual use (AU) differed significantly across disciplines,  $F(3, 252) = 5.694$ ,  $p = .001$ ,  $\eta^2 = .063$ , indicating a moderate effect size.

Post hoc Tukey HSD tests showed that students in Natural Sciences reported significantly higher actual use compared to those in Social Sciences ( $p = .013$ ) and Formal Sciences ( $p = .002$ ), while Arts and Humanities also differed from Formal Sciences ( $p = .031$ ). The findings suggest that actual engagement with AI tools varies, likely reflecting discipline-specific needs and exposure. Therefore, targeted interventions to enhance AI usage in disciplines with lower engagement, such as Formal Sciences and Social Sciences, may be warranted (See Appendix C for details).

*Table 4: Differences in AI Attitudes and Acceptance Across Academic Disciplines*

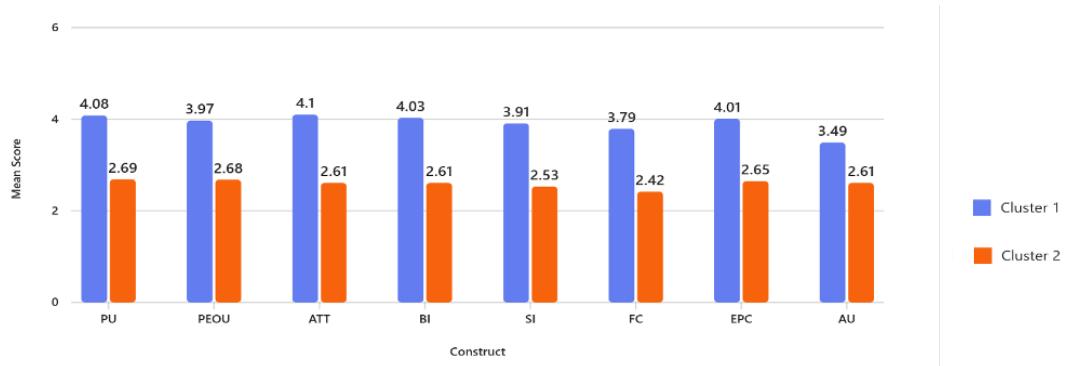
Variable	F	df	p	$\eta^2$
PU	0.837	(3, 252)	0.475	0.01
PEOU	1.647	(3, 252)	0.179	0.019
ATT	1.074	(3, 252)	0.361	0.013
BI	1.348	(3, 252)	0.259	0.016
SI	1.406	(3, 252)	0.241	0.016
FC	1.134	(3, 252)	0.336	0.013
EPC	0.848	(3, 252)	0.469	0.01
AU	5.694	(3, 252)	.001*	0.063

*Note: \* $p < .05$  indicates statistical significance*

### **3.3. Cluster Analysis of Student Attitudes and Acceptance of AI**

K-means cluster analysis was used to explore patterns in students' attitudes and acceptance of AI in higher education. The final cluster centers revealed that Cluster 1 ( $n = 144$ ) was characterized by higher mean scores across all constructs, as shown in figure 3. In contrast, Cluster 2 ( $n = 112$ ) exhibited lower mean scores on these constructs ( $PU = 2.69$ ,  $PEOU = 2.68$ ,  $ATT = 2.61$ ,  $BI = 2.61$ ,  $SI = 2.53$ ,  $FC = 2.42$ ,  $EPC = 2.65$ ,  $AU = 2.61$ ). ANOVA results indicated significant differences between clusters for all variables (all  $p < .001$ ), reflecting that the clustering procedure successfully distinguished between students with generally positive versus less positive attitudes and acceptance of AI. However, it should be noted that the F-tests are descriptive, as clusters were formed to maximize between-group differences, and significance levels are not corrected for this bias. Overall, the analysis suggests the presence of two distinct groups: one with high acceptance and positive attitudes toward AI, and another with lower acceptance and less favorable attitudes.

*Figure 3: Cluster Profiles of Students' Attitudes and Acceptance of AI across Constructs*



### 3.4. Key Determinants of Students' Intentions to Adopt AI Technologies

The hierarchical multiple regression analysis was performed to identify the key factors predicting students' behavioral intention to adopt AI technologies in Cambodian higher education. In Model 1, demographic variables (age, gender, discipline, and year of study) were entered and did not significantly predict behavioral intention,  $F(4, 249) = 1.12$ ,  $p = .35$ , explaining only 1.8% of the variance ( $R^2 = .018$ , adjusted  $R^2 = .002$ ).

*Table 5: Model Summary of Regression Analysis*

	R	R <sup>2</sup>	Adj. R <sup>2</sup>	SE	ΔF	df	p value
Model 1	0.133	0.018	0.002	0.977	1.117	4,249	0.349
Model 2	0.932	0.869	0.863	0.362	262.682	6,243	.000**

\*\* $p < .01$ .

When the psychological and perceptual variables, such as PU, PEOU, ATT, SI, FC, and EPC, were added in Model 2, the model significantly improved,  $\Delta F(6, 243) = 262.68$ ,  $p < .001$ , accounting for 86.9% of the variance in behavioral intention ( $R^2 = .869$ , adjusted  $R^2 = .863$ ).

*Table 6: Hierarchical Regression of Factors Predicting Students' Behavioral Intention*

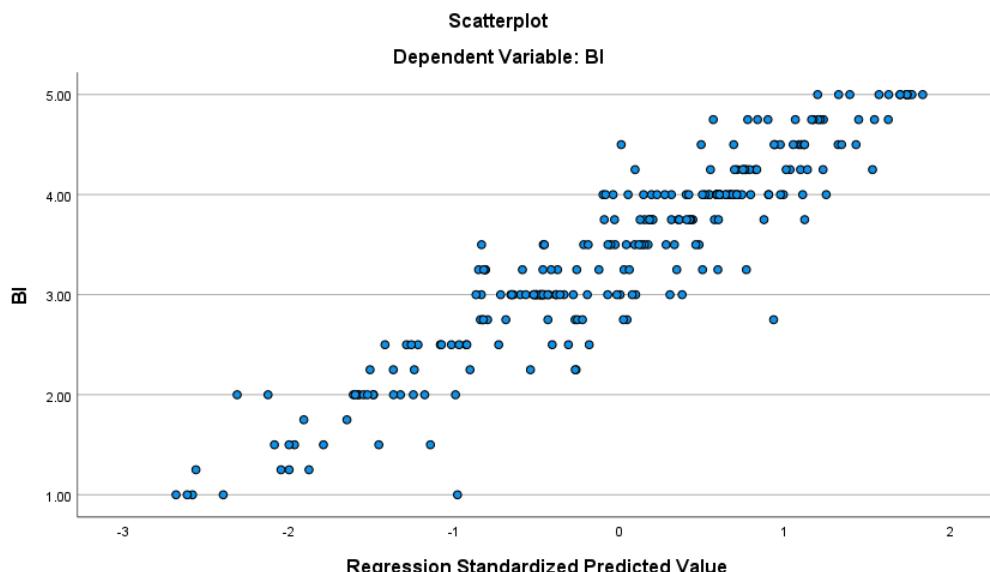
Model	Unstandardized Coefficients (B)		Std. Error	Standardized Coefficients (Beta)		t	Sig.
	Model 1	Model 2					
1	(Constant)	3.116	0.384			8.121	0.000
	Age	0.042	0.134	0.025	0.312	0.755	
	Gender	0.002	0.145	0.001	0.016	0.987	
	Discipline	0.002	0.058	0.002	0.03	0.976	
	Year	0.109	0.074	0.117	1.489	0.138	
	(Constant)	-0.109	0.172			-0.635	0.526
	Age	0.036	0.05	0.021	0.713	0.476	
	Gender	-0.022	0.054	-0.011	-0.414	0.68	
	Discipline	0.012	0.022	0.014	0.547	0.585	
	Year	0.015	0.028	0.016	0.556	0.579	
2	PU	0.055	0.049	0.054	1.132	0.259	
	PEOU	0.155	0.045	0.146	3.475	0.001	
	ATT	0.67	0.054	0.662	12.409	0.000	
	SI	0.005	0.056	0.005	0.089	0.929	
	FC	0.131	0.05	0.129	2.616	0.009	
	EPC	-0.007	0.045	-0.007	-0.15	0.881	

In the final model, three predictors made significant and unique contributions. Attitude toward AI emerged as the strongest predictor of behavioral intention ( $\beta = .662$ ,  $t = 12.41$ ,  $p < .001$ ), followed by PEOU ( $\beta = .146$ ,  $t = 3.48$ ,  $p = .001$ ) and FC ( $\beta = .129$ ,  $t = 2.62$ ,  $p = .009$ ). In contrast, PU, SI, and EPC were not significant predictors ( $p > .05$ ). The results suggest that students' positive attitudes toward AI, their perception of its ease of use, and the availability of supportive conditions play crucial roles in shaping their intention to adopt AI technologies in higher education.

Figure 4 presents a scatterplot illustrating the relationship between the regression standardized predicted values and the actual scores of students' behavioral intention (BI) to adopt AI technologies. The data points are closely aligned along a positive linear trend, indicating that the model's predicted values correspond well with the observed values of BI. This pattern suggests that the regression model fits the data appropriately and that the assumptions of linearity and homoscedasticity are reasonably met. In other words, higher predicted values are consistently associated with higher actual BI scores, confirming the model's strong predictive capability.

Overall, the findings emphasize that positive attitudes toward AI, perceived ease of use, and adequate institutional support are key determinants of students' behavioral intention to adopt AI technologies in Cambodian higher education settings.

*Figure 4: Scatterplot of Regression Standardized Predicted Values and Observed BI Scores*



## IV. Discussions

This chapter explored Cambodian university students' attitudes and acceptance of AI technologies, revealing several key findings that both align with and diverge from the broader literature on AI in education.

Consistent with the TAM and its extensions,<sup>22</sup> the present study found that attitude toward AI was the strongest predictor of students' behavioral intention to adopt AI, followed by PEOU and FC. This mirrors findings from recent studies in diverse contexts,<sup>23</sup> which also report that positive attitudes

## The Future of AI for Development

and usability are central to AI adoption among students. The strong influence of FC in the Cambodian context echoes calls in the literature for robust institutional infrastructure and technical support.<sup>24</sup>

However, while PU is often a significant predictor in TAM-based studies,<sup>25</sup> it was not a significant predictor in this study. This suggests that, in emerging contexts where AI is still novel, students may prioritize how easy and supported AI use feels over its perceived benefits, a nuance also noted by Chan & Hu (2023) and Baidoo-Anu et al. (2024), who observed that curiosity and usability often precede deep appreciation of usefulness.<sup>26</sup>

Contrary to some international findings where SI and ethical/privacy concerns (EPC) play a notable role,<sup>27</sup> this study found these factors did not significantly predict behavioral intention among Cambodian students. This may reflect Cambodia's proactive policy environment, where ethical guidelines and readiness assessments are already being integrated,<sup>28</sup> reducing student anxiety about these issues. It also suggests that, in this context, adoption decisions are more individually and institutionally driven than shaped by peer or leadership pressure, a finding that contrasts with studies in more collectivist or policy-mature settings.<sup>29</sup>

Despite generally positive attitudes and intentions, actual usage of AI tools remains moderate (mean scores 2.99–3.19 on a 5-point scale), with most students using AI "occasionally" to "often," but not "very often". This intention-behavior gap is widely reported in the literature,<sup>30</sup> and is often attributed to barriers such as limited access, digital literacy, and lack of localized resources.<sup>31</sup> The present findings reinforce the need for investments in infrastructure, professional development, and Khmer-language AI tools to bridge this gap.

The study found no significant gender differences in attitudes or acceptance of AI, and only discipline-specific differences in actual use, with Natural Science students reporting higher engagement. This aligns with recent multinational studies,<sup>32</sup> which also report minimal gender effects but highlight discipline as a moderating factor for actual AI engagement. Cluster analysis revealed two distinct groups: one with high acceptance and positive attitudes, and another with lower acceptance and less favorable attitudes. This pattern is echoed in the literature, suggesting that even within generally positive populations, substantial heterogeneity exists in readiness and enthusiasm for AI adoption.<sup>33</sup>

## V. Recommendations

Based on the findings of this chapter, some recommendations are proposed to strengthen the adoption and effective use of AI technologies in Cambodian higher education.

### **Prioritize the AI Training for Educators and Students**

- Institutions should invest in ongoing professional development that combines pedagogical upskilling with hands-on experience in AI tools. Training should be designed to various digital literacy levels, with special attention to creating Khmer-language and context-specific resources.

### **Enhance Institutional Infrastructure and Technical Support**

- HEIs should prioritize strengthening digital infrastructure, ensuring reliable internet connectivity, updated hardware, and access to educational AI applications. The technical support units should be established or expanded to offer real-time assistance, helping users overcome practical barriers.

### **Localize AI Tools and Content**

- There is an urgent need for AI tools, interfaces, and digital content that are adapted to the Cambodian socio-linguistic context. Key partnerships can foster the creation of Khmer-language AI resources that address local curriculum and learning preferences.
- *Foster a Positive, Inclusive Attitude Toward AI Institutions and the Ministries should develop outreach and awareness campaigns that use relatable success stories.*

### **Integrate AI Ethics Across Education Policy and Practice**

- While ethical concerns did not emerge as a current barrier, proactive inclusion of AI ethics, such as data privacy, bias prevention, and responsible use, should be mandatory in teacher training, student curricula, and policy statements.

### **Promote Equity and Inclusion in AI Adoption**

- Specific strategies are needed to ensure rural and under-resourced HEIs are not left behind. This includes providing subsidies, resource-sharing platforms, and piloting AI projects in diverse educational settings.

### **Encourage Collaborative Policymaking and International Partnerships**

- Finally, continued collaboration with UNESCO, regional partners, and the private sector will support Cambodia in developing national frameworks for AI in education. Regular policy reviews, stakeholder consultations, and pilot programs can drive adaptive, evidence-based approaches to AI integration.

higher education. Drawing on the Technology Acceptance Model and related frameworks, the findings reveal that students generally hold positive perceptions of AI's usefulness and ease of use, with attitude toward AI emerging as the strongest predictor of behavioral intention to adopt AI technologies. PEOU and FC also significantly contribute to students' intentions, highlighting the importance of user-friendly tools and robust institutional support.

Despite these positive attitudes and intentions, actual usage of AI tools remains moderate, suggesting a gap between willingness and practical engagement. This gap may be attributed to infrastructural limitations, varying levels of digital literacy, and the need for more localized and contextually relevant AI resources. Notably, neither gender nor academic discipline significantly influenced students' attitudes or acceptance of AI, except for actual use, which was higher among students in Natural Sciences.

The results highlight the need for targeted strategies to bridge the intention-usage gap. Investments in digital infrastructure, professional development, and the creation of Khmer-language AI resources are essential to foster more widespread and effective adoption. Additionally, while EPC did not significantly deter behavioral intention, ongoing efforts to integrate AI ethics into policy and practice remain important for responsible and inclusive digital transformation. In summary, the successful integration of AI in Cambodian higher education depends not only on technological advancement but also on fostering positive attitudes, enhancing usability, and ensuring equitable access and support.

## **VI. Conclusion**

This chapter provides a comprehensive examination of Cambodian university students' attitudes and acceptance of AI in

## References

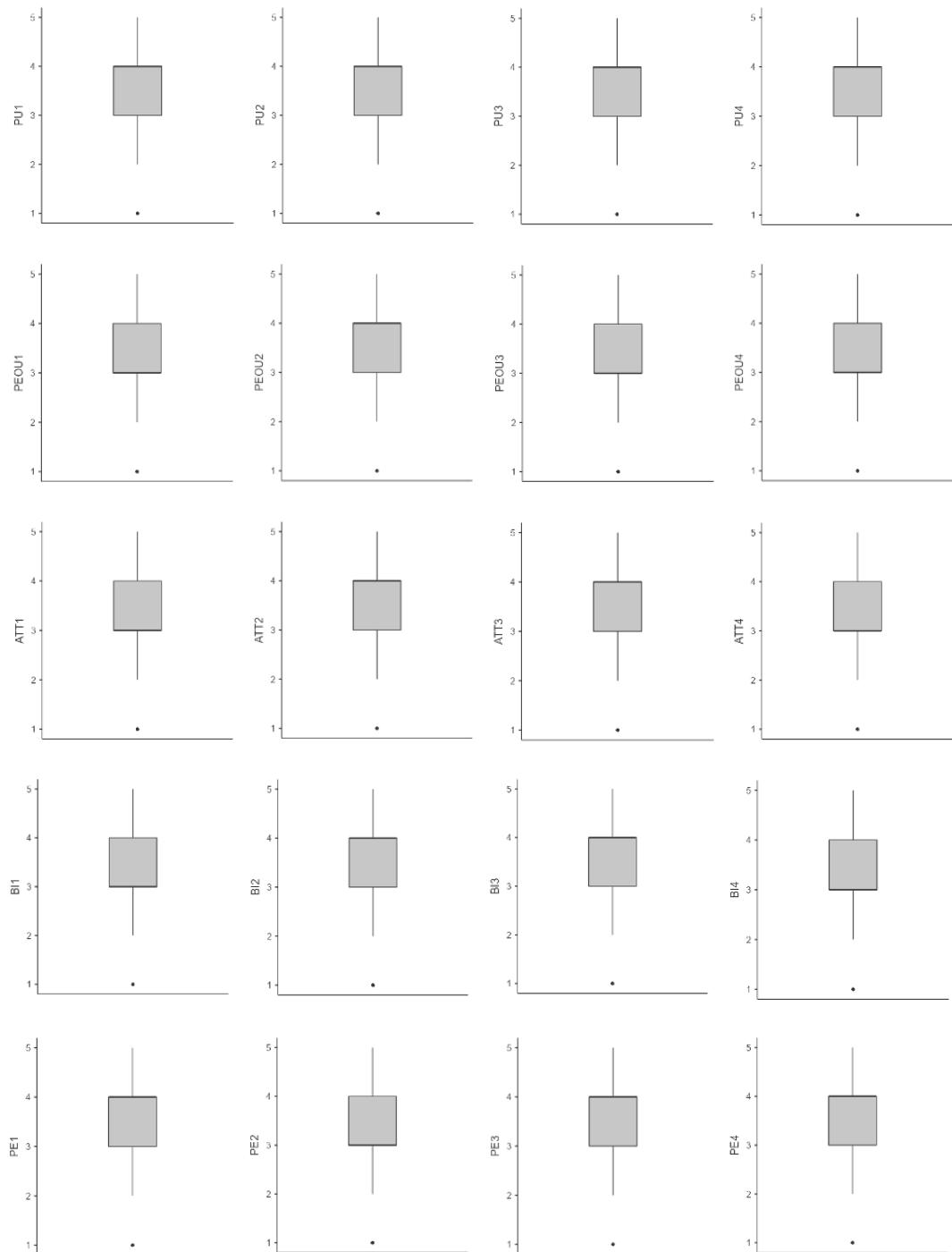
1. Wayne Holmes et al., "Ethics of AI in Education: Towards a Community-Wide Framework," *International Journal of Artificial Intelligence in Education* 32, no. 3 (2022): 504–26, <https://doi.org/10.1007/S40593-021-00239-1>.
2. Olaf Zawacki-Richter et al., "Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where Are the Educators?," *International Journal of Educational Technology in Higher Education* 16, no. 1 (2019), <https://doi.org/10.1186/S41239-019-0171-0>; Stanislav Ivanov et al., "Drivers of Generative AI Adoption in Higher Education through the Lens of the Theory of Planned Behavior," *Technology in Society* 77 (June 2024), <https://doi.org/10.1016/j.techsoc.2024.102521>.
3. Davis et al. (1989), n.d.
4. Timothy Teo, "Factors Influencing Teachers' Intention to Use Technology: Model Development and Test," *Computers and Education* 57, no. 4 (2011): 2432–40, <https://doi.org/10.1016/j.compedu.2011.06.008>.
5. Viswanath Venkatesh et al., "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* 27, no. 3 (2003): 425–78, <https://doi.org/10.2307/30036540>.
6. Heather Holden and Roy Rada, "Understanding the Influence of Perceived Usability and Technology Self-Efficacy on Teachers' Technology Acceptance," *Journal of Research on Technology in Education*, Routledge, June 1, 2011, world, <https://www.tandfonline.com/doi/abs/10.1080/15391523.2011.10782576>; Ivanov et al., "Drivers of Generative AI Adoption in Higher Education through the Lens of the Theory of Planned Behavior."
7. Jacques Bulchand-Gidumal, "Impact of Artificial Intelligence in Travel, Tourism, and Hospitality," in *Handbook of E-Tourism* (Springer, Cham, 2020), [https://doi.org/10.1007/978-3-030-05324-6\\_110-1](https://doi.org/10.1007/978-3-030-05324-6_110-1).
8. "Draft National Artificial Intelligence Strategy 2025-2030 (Version 5) - OD Mekong Datahub," accessed August 3, 2025, [https://data.opendevelopmentcambodia.net/library\\_record/draft-national-artificial-intelligence-strategy-2025-2030-version-5](https://data.opendevelopmentcambodia.net/library_record/draft-national-artificial-intelligence-strategy-2025-2030-version-5).
9. "Announcement on the Progress of Draft National Artificial Intelligence Strategy and the Opening of Public Consultation," Rajah & Tann Asia, n.d., accessed August 3, 2025, <https://www.rajahtannasia.com/viewpoints/announcement-on-the-progress-of-draft-national-artificial-intelligence-strategy-and-the-opening-of-public-consultation/>.
10. CADT, "Strategic Consultation on Validating Preliminary Results and Discussing Recommendations of the UNESCO's Ethics of AI Readiness Assessment in Cambodia," CADT, February 8, 2025, <https://cadt.edu.kh/workshop-on-validating-preliminary-results-and-discussing-recommendations-of-the-unescos-ethics-of-ai-readiness-assessment-in-cambodia/>.
11. CADT, "Strategic Consultation on Validating Preliminary Results and Discussing Recommendations of the UNESCO's Ethics of AI Readiness Assessment in Cambodia."
12. Phyrom Eam and Sopheak Song, *TECHNOLOGY INTEGRATION OF CAMBODIAN HIGHER EDUCATION: EMERGING CHANGES SINCE COVID-19 AND NOTABLE CHALLENGES* Paper Prepared for the Global Education Monitoring Report 2022 GEM Report Fellowship, 2023.
13. Selwyn et al., "What's next for Ed-Tech?"
14. Khmertime, [khmertimeskh.com/501597302/cambodias-ai-journey-overcoming-challenges-to-harness-advanced-tech/](https://khmertimeskh.com/501597302/cambodias-ai-journey-overcoming-challenges-to-harness-advanced-tech/).

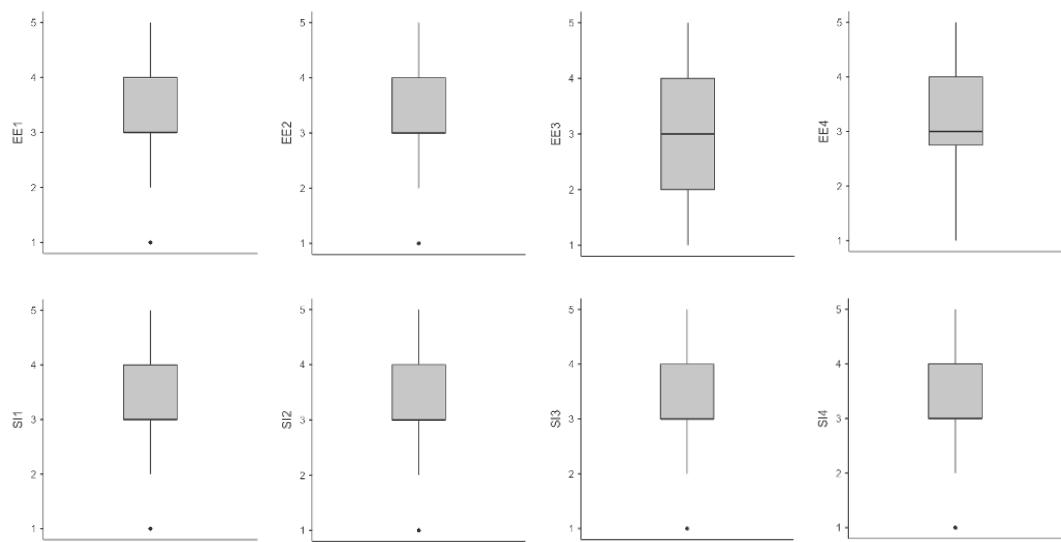
15. Edman Flores and Udam Mean, "Digital Transformation in Cambodian Higher Education: Current Trends and Future Directions," *Edelweiss Applied Science and Technology* 2025 (forthcoming), accessed August 4, 2025, <https://learning-gate.com/index.php/2576-8484/article/view/4267>.
16. UNESCO, "Cambodia Launches AI Readiness Assessment Report to Guide Ethical and Inclusive Digital Transformation," July 2025, <https://www.unesco.org/en/articles/cambodia-launches-ai-readiness-assessment-report-guide-ethical-and-inclusive-digital-transformation>.
17. John W. Creswell and Vicki L. Plano Clark, *Designing and Conducting Mixed Methods Research Approach* (2017).
18. Davis et al. (1989).
19. Venkatesh et al., "User Acceptance of Information Technology."
20. Field, A.P. (2018) *Discovering Statistics Using IBM SPSS Statistics*. 5th Edition, Sage, Newbury Park.
21. George, D., & Mallery, P. (2019). *IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference* (16th ed.). Routledge. <https://doi.org/10.4324/9780429056765>.
22. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems*, 27(3), 425–478. <https://doi.org/10.2307/30036540>; Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>.
23. Haq, M. Z. U., Cao, G., & Abukhait, R. M. Y. (2025). Examining students' attitudes and intentions towards using ChatGPT in higher education. *British Journal of Educational Technology*. <https://doi.org/10.1111/BJET.13582>; Tang, X., Yuan, Z., & Qu, S. (2025). Factors Influencing University Students' Behavioral Intention to Use Generative Artificial Intelligence for Educational Purposes Based on a Revised UTAUT2 Model. *Journal of Computer Assisted Learning*, 41(1). <https://doi.org/10.1111/JCAL.13105>.
24. Sok, S., & Heng, K. (2024a). Opportunities, challenges, and strategies for using ChatGPT in higher education: A literature review. *Journal of Digital Educational Technology*, 4(1), ep2401. <https://doi.org/10.30935/JDET/14027>; Mexhuani, B. (2024). Adopting Digital Tools in Higher Education: Opportunities, Challenges and Theoretical Insights. *European Journal of Education*. <https://doi.org/10.1111/EJED.12819>.
25. Kong, S. C., Zhu, J., & Yang, Y. N. (2025). Developing and validating a scale of empowerment in using artificial intelligence for problem-solving for senior secondary and university students. *Computers and Education: Artificial Intelligence*, 8, 100359. <https://doi.org/10.1016/j.CAEAI.2024.100359>; Abdaljaleel, et al. (2024). A multinational study on the factors influencing university students' attitudes and usage of ChatGPT. *Scientific Reports*, 14(1). <https://doi.org/10.1038/S41598-024-52549-8>.
26. Baidoo-Anu, D., et. al. (2024). Exploring student perspectives on generative artificial intelligence in higher education learning. *Discover Education*, 3(1). <https://doi.org/10.1007/S44217-024-00173-Z>; Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 1–18. <https://doi.org/10.1186/s41239-023-00411-8>.

27. Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). *Chatting and cheating: Ensuring academic integrity in the era of ChatGPT*. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>;
28. Akgun, S., & Greenhow, C. (2022). *Artificial intelligence in education: Addressing ethical challenges in K-12 settings*. *AI and Ethics*, 2(3), 431–440. <https://doi.org/10.1007/S43681-021-00096-7>.
29. UNESCO, "Cambodia Launches AI Readiness Assessment Report to Guide Ethical and Inclusive Digital Transformation," July 2025, <https://www.unesco.org/en/articles/cambodia-launches-ai-readiness-assessment-report-guide-ethical-and-inclusive-digital-transformation>.
30. Hew, K. F., Huang, W., Du, J., & Jia, C. (2023). *Using chatbots to support student goal setting and social presence in fully online activities: Learner engagement and perceptions*. *Journal of Computing in Higher Education*, 35(1), 40–68. <https://doi.org/10.1007/s12528-022-09338-x>;
31. Hoi, V. N., & Mu, G. M. (2021). *Perceived teacher support and students' acceptance of mobile-assisted language learning: Evidence from Vietnamese higher education context*. *British Journal of Educational Technology*, 52(2), 879–898. <https://doi.org/10.1111/BJET.13044>.
32. Kong, S. C., Zhu, J., & Yang, Y. N. (2025). *Developing and validating a scale of empowerment in using artificial intelligence for problem-solving for senior secondary and university students*. *Computers and Education: Artificial Intelligence*, 8, 100359. <https://doi.org/10.1016/J.CAEAI.2024.100359>.
33. Sok, S., Heng, K., & Pum, M. (2025). *Investigating High School Students' Attitudes Toward the Use of AI in Education: Evidence from Cambodia*. *SAGE Open*, 15(3). <https://doi.org/10.1177/21582440251353575>;
34. Fauzi, F., Tuhuteru, L., Sampe, F., Ausat, A. M. A., & Hatta, H. R. (2023). *Analysing the Role of ChatGPT in Improving Student Productivity in Higher Education*. *Journal on Education*, 5(4), 14886–14891. <https://doi.org/10.31004/JOE.V5I4.2563>.
35. Sok, S., & Heng, K. (2024a). *Opportunities, challenges, and strategies for using ChatGPT in higher education: A literature review*. *Journal of Digital Educational Technology*, 4(1), ep2401. <https://doi.org/10.30935/JDET/14027>.
36. Labrague, L. J., Aguilar-Rosales, R., Yboa, B. C., Sabio, J. B., & de los Santos, J. A. (2023). *Student nurses' attitudes, perceived utilization, and intention to adopt artificial intelligence (AI) technology in nursing practice: a cross-sectional study*. *Nurse Educ Pract*, 73, 103815. <https://doi.org/10.1016/j.npr.2023.103815>;
37. Abdaljaleel, et al. (2024). *A multinational study on the factors influencing university students' attitudes and usage of ChatGPT*. *Scientific Reports*, 14(1). <https://doi.org/10.1038/S41598-024-52549-8>.
38. Kong, S. C., Zhu, J., & Yang, Y. N. (2025). *Developing and validating a scale of empowerment in using artificial intelligence for problem-solving for senior secondary and university students*. *Computers and Education: Artificial Intelligence*, 8, 100359. <https://doi.org/10.1016/J.CAEAI.2024.100359>.

## Appendix

### 1. Boxplot for All Survey Items





## 2. Frequency of AI Tool Usage Across Academic Activities

Frequency of AI Tool Usage Across Academic Activities in Cambodian Higher Education (N=256)

Items	Mean	SD	1	2	3	4	5
How frequently do you use AI-based tools in your personal life? (AU1)	3.14	0.97	4.3 %	21.1 %	37.9 %	30.1 %	6.6 %
How often do you use AI-based tools specifically for educational purposes? (AU2)	3.18	0.96	3.1 %	21.9 %	35.2 %	33.2 %	6.6 %
How frequently do you use AI-based tools for professional or academic work outside of direct teaching/learning activities? (AU3)	3.15	0.96	3.5 %	21.1 %	39.5 %	28.5 %	7.4 %
How often do you use AI-based tools for the design and preparation of teaching/learning materials? (AU4)	3.06	0.99	6.3 %	20.7 %	40.6 %	25.8 %	6.6 %

How often do you use AI-based tools during direct teaching/learning activities, such as lectures, assignments, or projects? (AU5)	3.15	1.03	5.9 %	19.1 %	38.7 %	26.6 %	9.8 %
How often do you use AI-based tools specifically to teach, learn, or explore topics related to AI? (AU6)	3.19	0.99	4.7 %	19.5 %	35.2 %	33.2 %	7.4 %
How often do you use AI-based tools for analyzing or predicting performance, such as tracking progress or success in a course? (AU7)	3.01	1.08	9.0 %	21.5 %	38.3 %	21.9 %	9.4 %
How often do you use AI-based tools as intelligent tutoring systems (ITS) for additional guidance or support? (AU8)	3.06	1.06	8.6 %	19.9 %	36.3 %	27.3 %	7.8 %
How frequently do you use AI-based tools for exam preparation, assessments, or evaluations? (AU9)	2.99	1.11	10.5 %	21.9 %	34.0 %	25.4 %	8.2 %
How often do you use AI-based tools for adaptive or personalized learning to match learning preferences or needs? (AU10)	3.11	1.04	6.6 %	19.9 %	38.3 %	26.2 %	9.0 %

Note: Likert Scale (1 = Never; 2 = Rarely; 3 = Occasionally; 4 = Often; 5 = Very Often)

### 3. ANOVA for Comparison among Different Disciplines

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
PU	Between Groups	2.277	3	.759	.837	.475
	Within Groups	228.657	252	.907		
	Total	230.934	255			
PEOU	Between Groups	4.184	3	1.395	1.647	.179
	Within Groups	213.375	252	.847		
	Total	217.559	255			
ATT	Between Groups	3.009	3	1.003	1.074	.361
	Within Groups	235.456	252	.934		
	Total	238.465	255			

BI	Between Groups	3.837	3	1.279	1.348	.259
	Within Groups	239.148	252	.949		
	Total	242.984	255			
SI	Between Groups	3.599	3	1.200	1.406	.241
	Within Groups	214.975	252	.853		
	Total	218.574	255			
FC	Between Groups	3.130	3	1.043	1.134	.336
	Within Groups	231.899	252	.920		
	Total	235.029	255			
EPC	Between Groups	2.295	3	.765	.848	.469
	Within Groups	227.232	252	.902		
	Total	229.527	255			
AU	Between Groups	11.284	3	3.761	5.694	.001
	Within Groups	166.481	252	.661		
	Total	177.765	255			

ANOVA Effect Sizes <sup>a,b</sup>						
			Point Estimate	95% Confidence Interval		
				Lower	Upper	
PU	Eta-squared		.010	.000	.036	
	Epsilon-squared		-.002	-.012	.024	
	Omega-squared Fixed-effect		-.002	-.012	.024	
	Omega-squared Random-effect		-.001	-.004	.008	
PEOU	Eta-squared		.019	.000	.055	
	Epsilon-squared		.008	-.012	.043	
	Omega-squared Fixed-effect		.008	-.012	.043	
	Omega-squared Random-effect		.003	-.004	.015	
ATT	Eta-squared		.013	.000	.042	
	Epsilon-squared		.001	-.012	.030	
	Omega-squared Fixed-effect		.001	-.012	.030	
	Omega-squared Random-effect		.000	-.004	.010	
BI	Eta-squared		.016	.000	.048	
	Epsilon-squared		.004	-.012	.037	
	Omega-squared Fixed-effect		.004	-.012	.037	
	Omega-squared Random-effect		.001	-.004	.013	
SI	Eta-squared		.016	.000	.050	
	Epsilon-squared		.005	-.012	.038	
	Omega-squared Fixed-effect		.005	-.012	.038	
	Omega-squared Random-effect		.002	-.004	.013	

FC	Eta-squared	.013	.000	.043
	Epsilon-squared	.002	-.012	.032
	Omega-squared Fixed-effect	.002	-.012	.032
	Omega-squared Random-effect	.001	-.004	.011
EPC	Eta-squared	.010	.000	.036
	Epsilon-squared	-.002	-.012	.025
	Omega-squared Fixed-effect	-.002	-.012	.025
	Omega-squared Random-effect	-.001	-.004	.008
AU	Eta-squared	.063	.012	.120
	Epsilon-squared	.052	.000	.110
	Omega-squared Fixed-effect	.052	.000	.110
	Omega-squared Random-effect	.018	.000	.039

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

b. Negative but less biased estimates are retained, not rounded to zero.

### Multiple Comparisons

#### Tukey HSD

DV	(I) Discipline	(J) Discipline	Mean Difference		
			(I-J)	Std. Error	Sig.
PU	Social Science	Arts and Humanities	-.04263	.15329	.992
		Natural Science	.24583	.34564	.893
		Formal Science	-.21341	.16054	.545
	Arts and Humanities	Social Science	.04263	.15329	.992
		Natural Science	.28846	.36176	.856
		Formal Science	-.17078	.19281	.812
	Natural Science	Social Science	-.24583	.34564	.893
		Arts and Humanities	-.28846	.36176	.856
		Formal Science	-.45924	.36489	.590
	Formal Science	Social Science	.21341	.16054	.545
		Arts and Humanities	.17078	.19281	.812
		Natural Science	.45924	.36489	.590

PEOU	Social Science	Arts and Humanities	.21500	.14808	.468
		Natural Science	.59000	.33389	.292
		Formal Science	.00304	.15509	1.000
	Arts and Humanities	Social Science	-.21500	.14808	.468
		Natural Science	.37500	.34946	.706
		Formal Science	-.21196	.18625	.666
	Natural Science	Social Science	-.59000	.33389	.292
		Arts and Humanities	-.37500	.34946	.706
		Formal Science	-.58696	.35249	.344
	Formal Science	Social Science	-.00304	.15509	1.000
		Arts and Humanities	.21196	.18625	.666
		Natural Science	.58696	.35249	.344
ATT	Social Science	Arts and Humanities	.15564	.15555	.749
		Natural Science	.30708	.35075	.818
		Formal Science	-.15080	.16291	.791
	Arts and Humanities	Social Science	-.15564	.15555	.749
		Natural Science	.15144	.36710	.976
		Formal Science	-.30644	.19565	.400
	Natural Science	Social Science	-.30708	.35075	.818
		Arts and Humanities	-.15144	.36710	.976
		Formal Science	-.45788	.37028	.604
	Formal Science	Social Science	.15080	.16291	.791
		Arts and Humanities	.30644	.19565	.400
		Natural Science	.45788	.37028	.604
BI	Social Science	Arts and Humanities	.13468	.15677	.826
		Natural Science	.44958	.35348	.582
		Formal Science	-.16862	.16419	.734
	Arts and Humanities	Social Science	-.13468	.15677	.826
		Natural Science	.31490	.36997	.830
		Formal Science	-.30330	.19718	.416
	Natural Science	Social Science	-.44958	.35348	.582
		Arts and Humanities	-.31490	.36997	.830
		Formal Science	-.61821	.37317	.349
	Formal Science	Social Science	.16862	.16419	.734
		Arts and Humanities	.30330	.19718	.416
		Natural Science	.61821	.37317	.349

SI	Social Science	Arts and Humanities	-.05686	.14864	.981
		Natural Science	.65708	.33514	.206
		Formal Science	.00355	.15567	1.000
	Arts and Humanities	Social Science	.05686	.14864	.981
		Natural Science	.71394	.35077	.178
		Formal Science	.06041	.18695	.988
	Natural Science	Social Science	-.65708	.33514	.206
		Arts and Humanities	-.71394	.35077	.178
		Formal Science	-.65353	.35381	.254
	Formal Science	Social Science	-.00355	.15567	1.000
		Arts and Humanities	-.06041	.18695	.988
		Natural Science	.65353	.35381	.254
FC	Social Science	Arts and Humanities	.00865	.15438	1.000
		Natural Science	.63125	.34809	.269
		Formal Science	.07283	.16168	.969
	Arts and Humanities	Social Science	-.00865	.15438	1.000
		Natural Science	.62260	.36432	.321
		Formal Science	.06417	.19417	.988
	Natural Science	Social Science	-.63125	.34809	.269
		Arts and Humanities	-.62260	.36432	.321
		Formal Science	-.55842	.36747	.427
	Formal Science	Social Science	-.07283	.16168	.969
		Arts and Humanities	-.06417	.19417	.988
		Natural Science	.55842	.36747	.427
EPC	Social Science	Arts and Humanities	.03712	.15281	.995
		Natural Science	.49625	.34457	.475
		Formal Science	.13348	.16004	.838
	Arts and Humanities	Social Science	-.03712	.15281	.995
		Natural Science	.45913	.36063	.581
		Formal Science	.09636	.19221	.959
	Natural Science	Social Science	-.49625	.34457	.475
		Arts and Humanities	-.45913	.36063	.581
		Formal Science	-.36277	.36375	.751
	Formal Science	Social Science	-.13348	.16004	.838
		Arts and Humanities	-.09636	.19221	.959
		Natural Science	.36277	.36375	.751

AU	Social Science	Arts and Humanities	.23164	.13080	.290
		Natural Science	.90183*	.29493	.013
		Formal Science	-.22371	.13699	.362
	Arts and Humanities	Social Science	-.23164	.13080	.290
		Natural Science	.67019	.30868	.134
		Formal Science	-.45535*	.16452	.031
	Natural Science	Social Science	-.90183*	.29493	.013
		Arts and Humanities	-.67019	.30868	.134
		Formal Science	-1.12554*	.31135	.002
	Formal Science	Social Science	.22371	.13699	.362
		Arts and Humanities	.45535*	.16452	.031
		Natural Science	1.12554*	.31135	.002

\*The mean difference is significant at the 0.05 level.

#### 4. Cluster Analysis using K-Means

Cluster		
	1	2
PU	4.08	PU
PEOU	3.97	PEOU
ATT	4.10	ATT
BI	4.03	BI
SI	3.91	SI
FC	3.79	FC
EPC	4.01	EPC
AU	3.49	AU

ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
PU	122.746	1	.426	254	288.180	.000
PEOU	104.304	1	.446	254	233.925	.000
ATT	138.951	1	.392	254	354.662	.000
BI	126.748	1	.458	254	276.969	.000
SI	120.186	1	.387	254	310.274	.000
FC	116.928	1	.465	254	251.477	.000
EPC	116.885	1	.443	254	263.569	.000
AU	47.983	1	.511	254	93.909	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Number of Cases in each Cluster		
Cluster	1	144.000
	2	112.000
Valid		256.000
Missing		.000



Mr. Borey Be

*Borey Be is an independent researcher and Lecturer at various universities in Cambodia. He is currently a PhD candidate in educational science at the University of Cambodia. He holds a master's degree in educational leadership and policy (Comparative Education) from Beijing Normal University, China. His research centers on teaching and teacher education, special education, online teaching and learning, educational technology, and parental involvement. He has authored and co-authored several peer-reviewed publications and conference papers.*

# AI Meets Public Data: A Prompt-Based Agent for Intuitive Analysis of Government Datasets



Dr. Rina Buoy, Dr. Nguonly Taing, and H.E. Dr. Marry Kong

## I. Introduction

The evolution of artificial intelligence (AI), data, and software is intertwined with one another. Through AI evolution, software has continuously transitioned from Software 1.0 (i.e., rule-based, expert systems) to Software 2.0 (i.e., machine-learning-based systems) to Software 3.0 (i.e., prompt-based agents).<sup>1</sup> The AI agents, powered by vision language models (VLMs), large language models (LLMs), agentic design (e.g., ReAct),<sup>2</sup> and standardized tool calling protocols, e.g., model context protocol (MCP),<sup>3</sup> can perform complex goal-oriented tasks, such as data analysis and manipulations, autonomously by planning, reasoning, and executing using an array of MCP-exposed tools. OpenAI's Codex,<sup>4</sup> Cursor,<sup>5</sup> and Github Copilot,<sup>6</sup> are examples of AI coding agents. On the other hand, FinAgent,<sup>7</sup> and FinRobot,<sup>8</sup> are examples of financial agents that can perform an array of tasks from document data analysis to advanced analytical capabilities.

Recognizing the strategic importance of AI and agents, the National Artificial Intelligence Strategy (NAIS) of Cambodia has been drafted to extract maximum benefits while managing risks.<sup>9</sup> The current draft of NAIS (4<sup>th</sup> draft) highlights the need for developing and deploying AI technologies, data, and sectoral adoption. Additionally, the Cambodia Digital Economy and Society Framework 2021-2035 clearly features the roles of (big) data in the digital economy.<sup>10</sup> Thus, the General Department of Digital Economy (GDDE) under

the Ministry of Economy and Finance (MEF) launched the DataEF initiative, that is an open data lake for disseminating public economic and financial data to researchers, data scientists, and policy scholars.<sup>11</sup> At the time of writing, DataEF features 602 public datasets in 11 distinct categories from 33 organizations in different formats, such as unstructured PDF, structured CSV, etc.

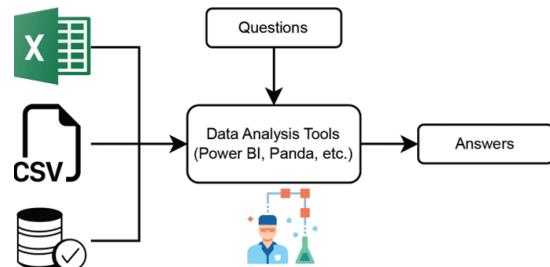
While having access to public data portals, such as DataEF, is of utmost importance, obtaining data-driven and actionable insights from such big data environments requires complex querying, manipulating, and analyzing techniques. These techniques require advanced data science expertise, a skill set that remains scarce in Cambodia.<sup>12</sup> *Thus, how can we remove the skill gap bottleneck to unlock data-driven insights and findings, particularly in the public sector?*

AI has come to our rescue. Broadly, AI refers to the science and engineering that make intelligent computer programs.<sup>13</sup> AI evolution has significantly transformed software from 1.0 (hard-coded rules) to 2.0 (data-driven) to 3.0 (prompt-based interface).<sup>14</sup> The idea of creating software through pure natural language instructions, once impossible, is now possible. Generative AI technologies, such as VLMs and LLMs, can understand human instructions in natural language or images and perform an array of complex tasks, such as coding and data analysis.<sup>15</sup>

Thus, in this chapter, we propose a prompt-based data analysis agent which can turn a natural language query into complex executable actions and codes agentically to arrive at an answer. Being agentic means that the system automates, reasons, and interacts with tools and environments.<sup>16</sup> Through a ChatGPT-like interface, a user can perform complex data analyses and manipulations on datasets housed in the DataEF platform. The qualitative experimental results have shown that it has the potential to reduce the time required for data analysis from days to minutes, thereby enhancing the reliability of information and significantly improving analytical productivity.

## II. The Proposed DataEF Agent

In this section, we introduce our proposed DataEF Agent. We begin by describing the conventional data analysis practice, as shown in Figure 1. Given user questions, a human data scientist needs to retrieve data from different sources (e.g., Excel, CSV, or SQL database). Various tools, such as PowerBI, Pandas (Python), can be used to manipulate and analyze the data to arrive at answers. Time-to-insight (Tol), which is defined as the time taken to arrive at answers, varies according to question complexity, skills, and technology stacks. Practically, Tol can take hours to days. Additionally, it is significantly prone to errors and miscommunication problems. For example, if the question is to analyze the country's trade deficits over the last five years by trading partners. First, the human data scientist must interpret the question, identify the user intents, and write SQL queries to select and filter the trade data over the last five years from the relevant databases. Next, the human data scientist can use either PowerBI or Pandas Dataframe in Python or other tools to compute trade deficits over time, visualize, and analyze the trends over time by trading partners.



*Figure 1: Conventional data analysis practice. A data scientist extracts data from different sources (Excel, CSV, or database) and uses various data analysis tools to answer various user questions*

In contrast, the architecture of the proposed DataEF agent is shown in Figure 2, leveraging VLMs and LLMs, the agentic patterns, and standard tool calling MCP servers. The main agent takes user questions, reasons, plans, writes code to execute, observes, and refines. Depending on the user's questions, the main agent decides, autonomously, to access the relevant datasets on DataEF, other data sources, and other AI models (e.g., machine translation or speech recognition) that are also exposed through MCP servers. The main agent can also delegate tasks to a specialized retrieval augmented generation (RAG),<sup>17</sup> agent to retrieve unstructured documents from vector databases. In case of failed executions, the main agent observes the errors and iteratively improves until the answers are reached. Depending on the user's questions, the agent can return the answers in text, audio, or image formats. Additionally, the conversations with the agent are saved in memory, allowing for follow-up questions at any time. There is no human intervention in the proposed architecture. For the same question of analyzing the country's trade deficits over the last five years, the main agent will, autonomously, reason and write code to query from the right datasets, manipulate the extracted data, and produce necessary graphs to achieve the goal.

With the proposed DataEF agent, it is now possible for users to ask questions in pure natural language or images through a ChatGPT-like interface, let the agent handle all complex processing steps, and get the answer in minutes. This allows quick iteration of data analysis tasks, which conventionally take both time and expertise to perform.

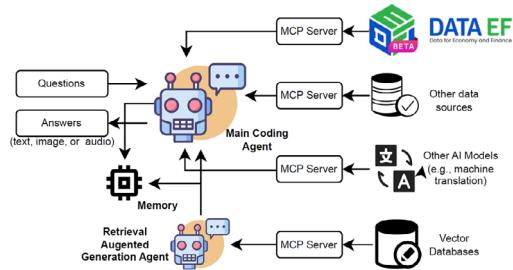


Figure 2: The Proposed DataEF Agent

### III. The Pilot Setup

In our pilot setup, we utilized the open-source, minimal agentic framework called ‘smolagents’.<sup>18</sup> As for the LLM, we utilized the OpenAI ‘o4-mini’ LLM model. We exposed DataEF to the main agent through the MCP servers. In addition to DataEF, we also allowed the agent to access the public data on the Cambodia Data Exchange (CamDX) Platform,<sup>19</sup> and Electronic Business Services (ESB),<sup>20</sup> portals also through the MCP servers. The RAG sub-agent is tasked with retrieving relevant unstructured documents (e.g., policy papers) which are stored in vector databases. Additionally, we allowed the main agent to access an in-house bidirectional Khmer-English machine translation model via MCP server so that the agent can translate questions or answers if required. Theoretically, we can expose as many AI models as we can to the main or sub agents. Thus, in this pilot setup, for a given question, the agent can perform the following tasks:

- Data analysis: The agent can reason, plan, and write complex queries to retrieve data from different data sources, and Python codes to manipulate and visualize.
- Machine translation: The agent can translate between Khmer and English whenever required.
- Knowledge retrieval: The agent, through its RAG sub-agent, can retrieve data from the knowledge base if required.
- Other data sources: Besides DataEF, the agent can also access the CamDX and ESB transaction databases.

### IV. Qualitative Evaluations

Based on the pilot setup mentioned above, we qualitatively evaluate the performance of our proposed DataEF agent by comparing with the human-created dashboards. Currently, there exist multiple data analytic dashboards that are manually crafted by human data scientists to answer some pre-defined data analysis questions related to the public economic and finance sector in DataEF. In total, there are 15 human-created dashboards comprising 81 charts based on various datasets, ranging from GDP growth to air quality data. For example, Figure 3 provides a human-created example of a gross domestic product (GDP) vs. time plot, highlighting the GDP trend. The actual time required to populate each data dashboard varies from several days to a few hours, depending on the user’s skills, familiarity with the data, and the complexity of the dashboard.

Considering the increasing number of available datasets on DataEF, it is not practical (if not impossible) to manually generate dashboards for all datasets. Thus, the proposed DataEF agent can augment the data scientists’ capabilities by giving them more time to think about good questions while the agent handles complex coding processes.

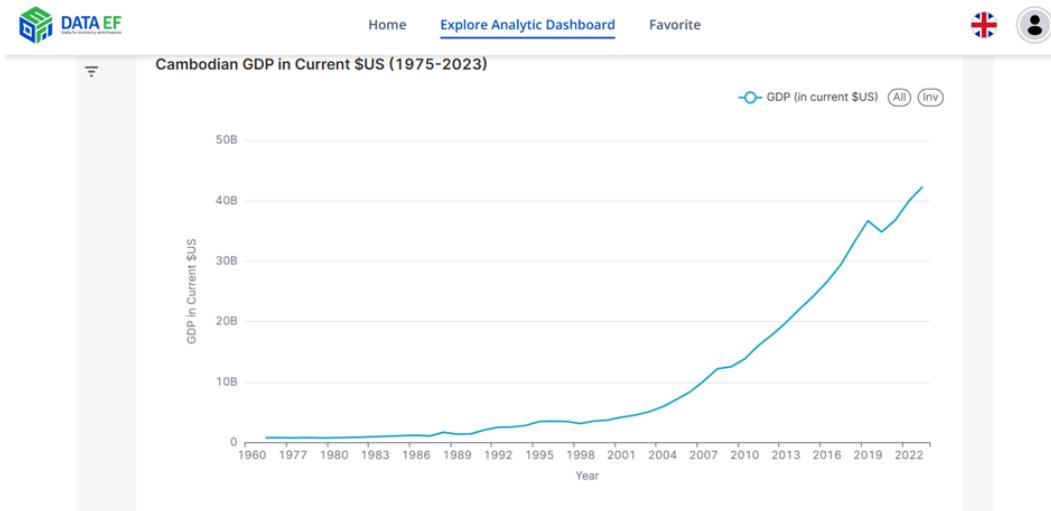


Figure 3: An example of human-created dashboards in DataEF

We prompted the proposed DataEF agent to generate the 81 charts and manually verified the outputs. In all cases, the DataEF agent successfully produced charts comparable to the human-created ones. Based on common query patterns and chart types in the DataEF platform, Figure 4, Figure 5, and Figure 6 provide some selected comparisons between human-generated versus agent-generated plots for qualitative evaluation. In all cases, we observed that the agent managed to reproduce the same analysis plots and arrive at the same conclusions as the human data scientists did. In all cases, the agents simply take the user's questions in natural language, as shown in the figures, reason, and write codes autonomously to achieve the goals. The number of iterations or steps the agent took varies according to the complexity of the questions. However, it was between two and five steps.

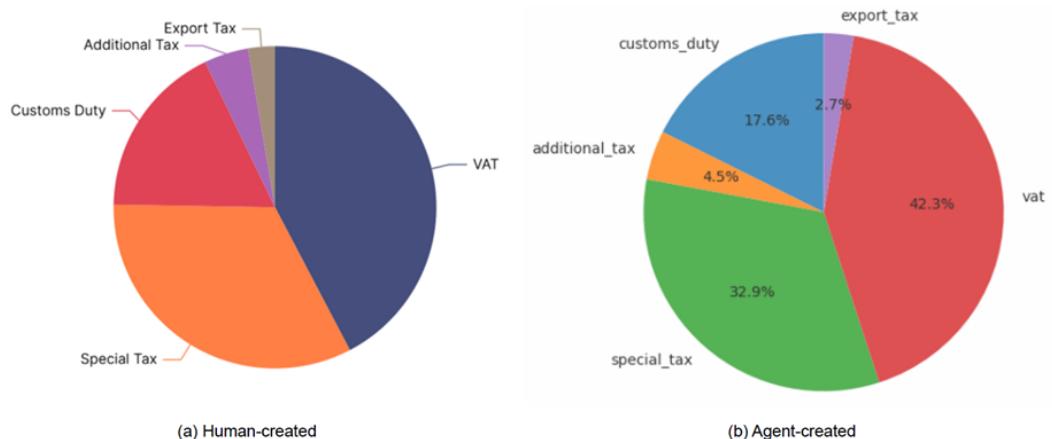


Figure 4: Case 1 - Human-generated (a) vs. agent-generated (b). Here is the user question (prompt) to the agent: 'Using revenue data collection, get the total revenue in 2024 by tax type, and plot a pie chart'

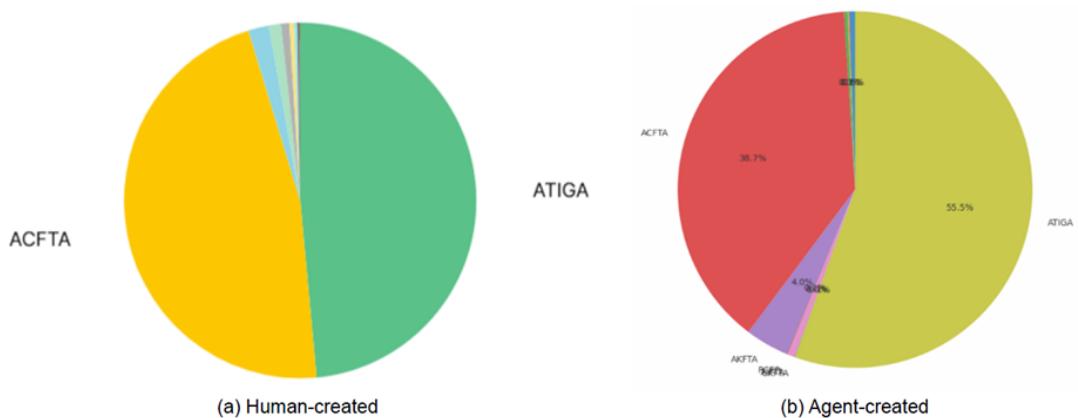


Figure 5: Case 2 - Human-generated (a) vs. agent-generated (b). Here is the user question (prompt) to the agent: 'Using revenue data collection, get the total revenue by free trade agreements in 2024 and plot a pie. Exclude the no agreement from the analysis'

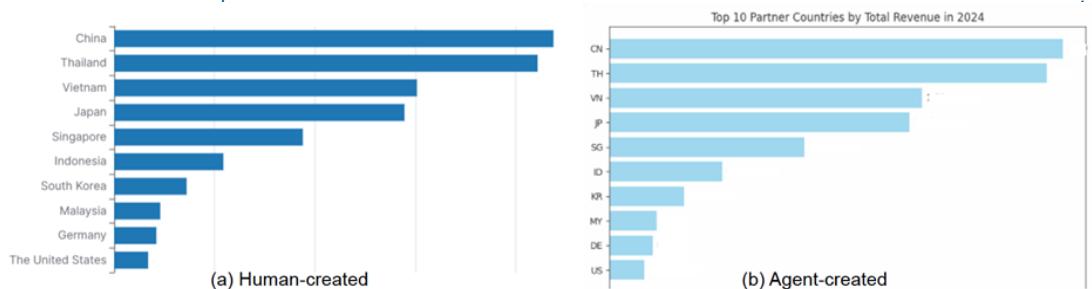


Figure 6: Case 3 - Human-generated (a) vs. agent-generated (b). Here is the user question (prompt) to the agent: 'Using revenue data collection, get the total revenue by 10 partner countries in 2024 and plot a horizontal bar'

Next, Figure 7, Figure 8, Figure 9, and Figure 10 provide the agent responses for different user prompts where we do not have human-generated answers to compare with. In Figure 7 and Figure 8, the agent was asked retrieval-based questions, where the agent needed to first retrieve some documents through its RAG sub-agent and synthesize the final answer. As shown in Figure 7, when the agent was being asked about a non-relevant question ('Who is Donald Trump?'), the agent correctly declined to answer and thus, prevented hallucinated responses. This is an expected behavior for a non-relevant question. However, as shown in Figure 7, when asked about a relevant question, the agent managed to answer and translate into Khmer as instructed.

When asked about the questions about CamDX and ESB, where the agent needs to access other data sources besides DataEF, the agent managed to write codes to query from the right data

**Step 4**

Thought: The question "Who is Donald Trump?" cannot be answered with the available translation tools, and no other tool is applicable. I'm sorry, but I can't help with that.

```
final_answer("I'm sorry, but I can't help with that.")
```

Used tool python\_interpreter

Execution Logs

**Step 4** | Input tokens: 2,670 | Output tokens: 589 | Duration: 11.36s

**Final answer:**  
I'm sorry, but I can't help with that.

Figure 7: Agent-generated response. Here is the user question (prompt) to the agent: 'Who is Donald Trump?'

Figure 8: Agent-generated response. Here is the user question (prompt) to the agent: 'Tell me more about MEF IT Architecture. Translate the final answer to Khmer'

Final answer:

## Top 5 ESB Member Systems by Approved Applications

System	Percentage of Approved Applications
National Social Security Fund	30.1%
Insurance Agent Registration as Natural Person	59.3%
Other (including Health Insurance, State Tax Office, and others)	0.6%

Figure 9: Agent-generated response. Here is the user question (prompt) to the agent: 'Identify the top electronic services for business in terms of approved applications? Plot a pie chart'

**Step 2** | Input tokens: 2,922 | Output tokens: 256 | Duration: 4.17s

---

**Final answer:** 429698

Figure 10: Agent-generated response. Here is the user question (prompt) to the agent: ‘Get total number of failed requests on CamDX in the last 3 months’

## V. Limitations and Future Work

The following limitations and future directions are identified as follows. Firstly, the proof-of-concept (PoC) pilot DataEF agent is based on the proprietary OpenAI 'o4-mini' model that may incur a significant API cost and is not suitable for handling sensitive data sources. Thus, future work should focus on experimentations with open-source LLM models or data anonymization approaches. Secondly, the evaluation is primarily qualitative and subjective. Thus, a more quantitative and scalable evaluation should be proposed. For example, using a VLM/LLM model to judge whether human-generated and agent-generated plots and analysis are equivalent. Lastly, the evaluation is limited in scale and scope because of limited evaluation data (i.e., limited number of human-generated dashboards). Thus, future work should focus on constructing a more comprehensive and open benchmark dataset.

Despite some limitations to overcome, the pilot DataEF agent has clearly demonstrated strong coding and analysis capabilities as human data scientists do. Thus, using the data analysis agent for routine data analysis tasks can significantly save processing time and thus result in a significant increase in analytical productivity, especially in the public sector. This does not imply that the agent can replace human data scientists. Instead, the agent amplifies the data scientists by allowing them more time to think about more important questions, which yields more important insights and findings. The successful PoC of this pilot DataEF agent highlights that instead of pre-defining the data analysis or plots in a given dashboard, we can, otherwise, dynamically and on demand, request specific analysis via a conversation agent.

## VI. Conclusion

In this chapter, we proposed a pilot DataEF agent that can perform complex data analysis on the datasets in DataEF and beyond as human data scientists do. Qualitative assessment suggests that using the data analysis agent can yield human-level performance, save time, and thus, significantly increase analytical productivity, especially in the public sector. Nonetheless, more rigorous and scalable evaluations and experimentations need to be done to further validate the proposed DataEF agent before full deployment within the ministry and the public.

## References

1. "Software 2.0", Andrej Karpathy, accessed July 22, 2025, <https://karpathy.medium.com/software-2-0-a64152b37c35>.
2. Yao Shunyu, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao, "ReAct: Synergizing Reasoning and Acting in Language Models", arXiv, 2024, arXiv.2210.03629.
3. "Introducing the Model Context Protocol", Anthropic, accessed July 22, 2025, <https://www.anthropic.com/news/model-context-protocol>.
4. "OpenAI Codex", OpenAI, accessed July 22, 2025, <https://openai.com/codex/>.
5. "Cursor: The AI Code Editor", Cursor, accessed July 22, 2025, <https://cursor.com/>.
6. "GitHub Copilot: Your AI Pair Programmer", GitHub, accessed July 22, 2025, <https://github.com/features/copilot>.
7. Zhang Wentao, Lingxuan Zhao, Haochong Xia, Shuo Sun, Jiaze Sun, Molei Qin, Xinyi Li, Yuqing Zhao, Yilei Zhao, Xinyu Cai, Longtao Zheng, Xinrun Wang, and Bo An, "A Multimodal Foundation Agent for Financial Trading: Tool-Augmented, Diversified, and Generalist", arXiv, 2024, arXiv.2402.18485.
8. Yang Hongyang, Boyu Zhang, Neng Wang, Cheng Guo, Xiaoli Zhang, Likun Lin, Junlin Wang, Tianyu Zhou, Mao Guan, Runjia Zhang, and Christina Dan Wang, "FinRobot: An Open-Source AI Agent Platform for Financial Applications Using Large Language Models", arXiv, 2024, arXiv.2405.14767.
9. Ministry of Post and Telecommunications, "National Artificial Intelligence Strategy 2025-2030" (Ministry of Post and Telecommunications, 2025).
10. Royal Government of Cambodia, "Cambodia Digital Economy and Society Policy Framework 2021-2035" (Royal Government of Cambodia, 2021).
11. "Data for Economic and Finance", General Department of Digital Economy, accessed July 22, 2025, <https://data.mef.gov.kh/about-us>.
12. "Cambodia Launches Its First Competency Framework on Digital, Media and Information Literacy to Empower Citizens in Today's Digital Society", UNESCO, accessed July 22, 2025, <https://www.unesco.org/en/articles/cambodia-launches-its-first-competency-framework-digital-media-and-information-literacy-empower>.
13. "What is AI?", McCarthy John, accessed July 22, 2025, <https://www-formal.stanford.edu/jmc/whatisai.pdf>.
14. "How AI Transforms Everyday Life", Y Combinator, accessed July 22, 2025, <https://www.youtube.com/watch?v=LCEmiRjPEtQ&t=604s>.
15. Mumuni Alhassan and Fuseini Mumuni, "Large language models for artificial general intelligence (AGI): A survey of foundational principles and approaches", arXiv, 2025, arXiv:2501.0315.
16. Sapkota Ranjan, Konstantinos I. Roumeliotis, and Manoj Karkee, "AI Agents vs. Agentic AI: A Conceptual Taxonomy, Applications and Challenges.", arXiv, 2025, arXiv.2505.10468.

17. *Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks.", arXiv, 2021, arXiv. 2005.11401.*
18. *Aymeric Roucher, Albert Villanova del Moral, Thomas Wolf, Leandro von Werra, and Erik Kaunismäki, "smolagents: A Smol Library to Build Great Agentic Systems", Github, 2025, <https://github.com/huggingface/smolagents>.*
19. *"CamDX Monitoring Tools", CamDX, accessed July 22, 2025, <https://monitoring.camdx.gov.kh/>.*
20. *"Statistics of business registration by E-Services", ESB, accessed July 22, 2025, <https://esb-monitoring.camdx.gov.kh/>.*



**Dr. Rina Buoy**

*Dr. Rina Buoy earned his B.S. and M.S. degrees in petroleum engineering from Universiti Teknologi PETRONAS, Malaysia, and an M.S. in IT engineering from the Royal University of Phnom Penh, Cambodia, both in 2022. He worked as an R&D team lead at Three60 Energy Asia from 2012 to 2019. Since 2020, he has been an applied AI researcher at Cambodia's Techo Startup Center of Ministry of Economy and Finance. In 2024, he completed his PhD from Osaka Metropolitan University, where he is now a visiting researcher. He is also a research fellow at AI Forum, Cambodia. Additionally, he is also an adjunct faculty member at American University of Phnom Penh and CamTech University.*



**Dr. Nguonly Taing**

*Dr. Taing Nguonly is a significant presence in Cambodia's digital economy sector, holding key roles such as Director General of the General Department of Digital Economy of the Ministry of Economy and Finance and Executive Director of Techo Startup Center. His responsibilities encompass overseeing diverse programs like the Cambodia Data Exchange (CamDX), Online Business Registration, Startup Development, Startup Cambodia National Program, and Enterprises Go Digital Program. Engaged in various research ventures spanning applied, policy, and AI research, Dr. Nguonly has been instrumental in developing crucial digital enabling platforms to advance Cambodia's digital economy.*

*His academic background features a B.Sc. in Computer Science from the Royal University of Phnom Penh, an M.Sc. in ICT from Waseda University, Japan, and a Ph.D. in Computer Science from Dresden University of Technology, Germany.*



**H.E. Dr. Marry Kong**

*H.E. Dr. Marry Kong is currently the Secretary of State of the Ministry of Economy and Finance, Chairman of the Board of Directors of Techo Startup Center, Director of the Entrepreneurship Development Fund, and Secretary General of the General Secretariat of the Digital Economy and Business Committee. His main roles at the Ministry involve driving reforms through the use of digital technology and platforms, developing startups and digital entrepreneurship and SMEs to go digital, and promoting innovation in financial technology (FinTech). Dr. Marry Kong obtained his Master's and Doctoral degrees in Telecommunications Technology and Information Technology from Waseda University, Japan.*

# Training the Trainers – Scaling AI Annotation Through Inclusive Workforce Development

---



*Songheng Tieng, Sokheng Pok, and Chantola Khet*

## I. Introduction

A huge amount of high-quality, structured data is a primary factor in developing global economic systems around AI. The datasets used to train advanced machine learning models are what make them work. These models can do things like drive cars, diagnose medical problems, and deliver personalized content.<sup>1</sup> Data annotation, the process of labeling raw data so that algorithms can understand it, has become a multi-billion-dollar global industry because of this data dependency. The process is not at all mechanical; it needs people to understand nuance, context, and ambiguity, which affects how AI systems will act in the end. However, as AI has grown at an incredible rate, it has hit a major roadblock because this work requires a lot of people. The problem is not just how many people there are, but also how to find and train a skilled and dependable workforce that can provide the consistently accurate and nuanced data that advanced AI systems need. Not doing so not only slows down innovation, but it also puts new technologies at risk of being built on top of societal biases and mistakes.<sup>2</sup>

The most common ways to get this kind of work done, which usually involve using fragmented crowdsourcing platforms or traditional business process outsourcing (BPO), have raised a lot of concerns about data quality, scalability, and the ethics of the workers.<sup>3</sup> These methods often see human input as a commodity, which can lead to worker instability, inconsistent quality, and a

lack of long-term skill building in communities.<sup>4</sup> For example, crowdsourcing is good for big projects, but it has trouble with tasks that need a lot of domain knowledge or understanding of the situation. Also, the fact that the workers are anonymous makes quality control a constant problem for managers. Traditional BPOs may have more control, but they also have a lot of turnover and a rigid structure that does not work well for the changing and flexible needs of AI development. There is a big gap in both academic and industry discussions because of the focus on technology and cutting costs instead of developing talent. There is a lot of research on algorithmic and platform-based solutions, but not as much on human-centered, scalable, and fair workforce development models that see workers as long-term assets to be developed. This mistake shows a strategic weakness in the AI supply chain, making it reliant on weak labor models that will not last in the long run.

This study fills in this gap in knowledge by looking at a different framework: an inclusive “train-the-trainer” model that aims to create long-lasting skill ecosystems in communities that do not have enough of them. This chapter is a case study of Digital Divide Data (DDD) in Cambodia, an organisation that has been championed impact sourcing, to answer the main research question: How can a “train-the-trainer” model, combined with soft skills and gender-equity interventions, effectively and inclusively scale high-quality AI annotation capacity?

This chapter makes the case that the AI industry's data bottleneck can be overcome by investing in local leadership and teaching skills that prepare workers to do more than just their jobs. They should also be able to train, mentor, and manage others. This method creates a self-sustaining cycle of knowledge transfer and quality assurance that builds a workforce that can grow and adapt. In this way, it creates big, long-term social and economic benefits that fit with the global development agenda and technological progress. The paper starts by looking at what others have written about digital labor and workforce development. Then it goes on to explain the case study method and analysis. Finally, it gives policymakers and tech companies ideas for how to make AI more responsible and actionable in the future.

## II. Literature Review

### 1. The Global AI Annotation Landscape

The market for data annotation, which is now an important part of the machine learning lifecycle, has grown and expanded in complexity along with the AI industry.<sup>5</sup> This "data work" includes a range of different tasks from basic tasks like sorting images and typing them up to very difficult ones like semantic segmentation of medical images or sentiment analysis of complex text. The quality of the data that people give to AI models has a direct and significant effect on how well they work. If the data is poorly labelled or biased, the algorithms will have bad and unfair results.<sup>6</sup> Crowdsourcing, which gives small tasks to people all over the world, is very fast and large.<sup>7</sup> But research has shown that it has many problems, such as problems with making sure the data is accurate, high worker turnover, and the fact that the model is not good for complex, multi-step annotation tasks that need a lot of focus and knowledge in the field.<sup>8</sup>

Traditional BPO models, which are often based in developing countries, can offer a more controlled setting and better quality. People often say that they are inflexible, slow to change, and that they copy the way things work on the factory floor, which limits workers' ability to learn new skills and be independent. Both models have been shown to create a group of "ghost workers" who do the invisible but necessary work that makes AI work. Gray and Suri say that these workers often do not have the social safety nets, chances to move up in their careers, or professional recognition that regular employees do. This makes them more vulnerable to algorithmic management and unstable wages.<sup>9</sup> This way of thinking sees human labor as a commodity that can be traded rather than as a source of expertise. This makes it hard to build a truly skilled and stable workforce. Le Ludec, Cornet, and Casilli make the case that the true impact of AI on labor is fostering an increasingly precarious globalised workforce of underpaid pieceworkers.<sup>10</sup>

**Spectrum of Workforce Models for AI Annotation**



### ***Impact Sourcing and Inclusive Workforce Development***

Impact sourcing is a direct response to the problems with these extractive labor models. As a business practice, it intentionally hires people from poor or underserved communities to do work that is outsourced.<sup>11</sup> The model is defined by its two goals: to give clients high-quality services and to give people who would otherwise have few options job and growth opportunities. Studies show that when done right, impact sourcing can help employees

make more money, learn new skills, and move up in society.<sup>12</sup> But for this model to work, there needs to be significant investment in workforce development that goes beyond just training people to do their jobs. The best programs teach both technical skills and important “soft skills” like communication, critical thinking, teamwork, and leadership that are necessary for long-term career success.<sup>13</sup> This all-encompassing view of the employee as a whole person with a career path is key to turning a job into a long-term source of income.

## 2. The “Train-the-Trainer” Methodology

The “train-the-trainer” (TtT) model is a key way to help more people learn new skills. This teaching method focuses on giving a small group of people the knowledge and skills they need to teach others. The TtT model comes from corporate training and community education. It acts as a force multiplier, making it possible for skills to spread quickly and in a decentralized way.<sup>14</sup> The model is based on social learning theory, which says that people can learn well by watching and copying what others do, especially peers who are credible.<sup>15</sup> In the workplace, it means making a plan to improve leadership and skills within the organization. An organization can create a learning ecosystem that is more adaptable and culturally relevant than one that relies on outside experts, who are often from other countries, by giving local employees the power to become mentors and trainers.<sup>16</sup> This internal capacity makes organizations more resilient by letting teams quickly adjust to new project needs and technologies without having to spend a lot of money on outside help. However, even though the TtT model has been shown to work well in other fields, there is still not much academic research on how it can be used as a formal strategy for building a large, high-quality, diverse workforce in the AI annotation sector.

## 3. The Cambodian Context

The contemporary social demographics, technology landscape, and economic context of Cambodia make for an ideal context to explore the “train the trainer” methodology. More than 60% of the people in the country are under 30 years old, so there is a relatively large population of young adults in the workforce.<sup>17</sup> This group of people, along with better digital infrastructure, shows that there is a lot of room for growth in the technology services sector. The Royal Government of Cambodia has seen this potential and started the “Cambodia Digital Economy and Society Policy Framework 2021–2035” to help the country become a lively digital economy and society.<sup>18</sup> At the same time, the country has ongoing problems, such as a gap between the skills taught in schools and what businesses need, and differences in opportunity between cities and rural areas, especially for women and people with disabilities.<sup>19</sup> In this situation, Cambodia is a very important test case to establish if inclusive workforce models can use the country’s demographic dividend to build a digital economy that is both competitive and fair, in line with the country’s stated policy. This study fills an important gap by looking at a specific, on-the-ground intervention that was made to meet a global technological challenge while also taking into account local social and economic needs.

## III. Methodology

This study uses a qualitative, single-case study research design that looks at secondary data. The case study method is great for looking into a modern phenomenon in depth in its real-world setting, especially when the lines between the phenomenon and the setting are not very clear.<sup>20</sup> Since the goal of this chapter is to learn how a certain workforce development model works, a case study of a pioneering organisation gives it the depth and focus it needs. The choice to only use secondary sources is a practical way to deal with research

limitations, and it sets the limits of the inquiry as an examination of the publicly stated model and its reported effects. We compare this single case to other models, based on recent research on crowdsourcing and BPO work models.

### **1. Case Selection**

This chapter chose DDD in Cambodia as a use case. An instrumental case is one that is chosen to shed light on a bigger issue or to improve a theoretical explanation. In this case, it is the effectiveness of the “train-the-trainer” model within an impact sourcing framework.<sup>21</sup> DDD presents a particularly strong case for this study for several key reasons. First, it is a well-known and well-respected pioneer of the impact sourcing model that has been in business for more than 20 years since it was founded in 2001. Second, its mission clearly targets young people, women, and people with disabilities who do not get enough help, which is in line with the paper’s focus on inclusive development. Third, DDD’s work in Cambodia is well-documented, though in different ways, by both its own publications and those of other groups. This gives us a lot of secondary data to work with.

### **2. Data Sources and Collection**

This study got the data for this study by carefully looking through publicly available documents that were made between 2010 and now. There were no interviews or surveys done to collect primary data. The body of secondary data includes annual reports, impact reports, white papers, and blog posts that are organisational publications available on the official DDD website, academic and grey literature such as research papers, conference papers, and working papers that look at or talk about DDD’s model, third-party reports from organisations, foundations, and partners (like the World Bank, Rockefeller Foundation, and Global Impact Sourcing Coalition) that have given money to or looked at DDD’s work, and media coverage in the form of articles and

features about DDD’s work and effects in Cambodia from well-known international and local news sources.

### **3. Analytical Framework**

This study used a thematic synthesis method to look at the data it had gathered. This required reviewing documents to find patterns and themes that kept coming up that were related to the research question.<sup>22</sup> The analysis was based on the main ideas found in the literature review: (1) putting the workforce development model into practice, with a focus on finding parts of the “train-the-trainer” approach; (2) combining soft skills training with technical skills training; (3) specific actions taken to make sure everyone is treated fairly and included; and (4) the reported effects on participants’ social and economic situations. This study puts together information from a variety of data sources to create a clear story about the DDD model and then compares it to the theoretical frameworks that were given.

## **IV. Findings: The Digital Divide Data Model in Cambodia**

A look at secondary sources shows that DDD Cambodia has a complex workforce development model that is very similar to the principles of impact sourcing and includes some parts of a “train-the-trainer” system.

### **1. The “Train-the-Trainer” Model in Practice**

DDD offers social mobility career paths aligned with the overarching approach of “train-the trainer”. According to DDD’s own reports, there is a tiered system in which new hires, or “operators,” learn basic data skills. After that, operators who do well are given jobs like Quality Assurance (QA) specialist, Project Coordinator, or Team Leader. In these higher-level jobs, they no longer do direct production. Instead, they review the work of junior operators, give them feedback, and train new hires. The Global Impact Sourcing Coalition’s

2019 report, which uses DDD as a case study, says that all of DDD's middle managers are promoted from within.<sup>23</sup> This one number clearly shows how a closed-loop system for developing talent works. Since 2001, DDD has hired and trained more than 8,000 young Cambodians, resulting in a large internal talent pipeline towards leadership positions.<sup>24</sup> It sets up a scalable, internal way to transfer skills, where experienced local youth who have mastered the technical and soft skills needed are in charge of building the capacity of the next group, which is the main idea behind TtT. Because the company invests in the person rather than just the role, a successful DDD employee is well-prepared for higher-skilled jobs beyond the company. The model helps build the technical skills and leadership of a new generation of Cambodian professionals by giving them direct experience that teaches them both job skills and how to be a leader.

## 2. Integration of Soft Skills for Leadership and Quality

DDD's publications all have a common theme: they stress the importance of holistic development. The structure of DDD's work-study program gives students 20 hours a week to work and 20 hours a week to learn and train.<sup>25</sup> This training time is not only for learning how to annotate data, but also for developing important soft skills. A report on developing skills in young people found that DDD's training includes more than 300 hours of instruction in English language skills, professional communication, critical thinking, and problem-solving.<sup>26</sup> This is done through a structured program that lasts for several years. According to an International Finance Corporation's case study, students spend their first year learning English and how to use computers. They get professional development training in their second and third years. This includes project management and leadership modules that are meant to help them get promoted.<sup>27</sup> There are two reasons given for this. First, these skills are thought

to be necessary for success in a career over the long term. Second, they are seen as the building blocks of leadership. DDD gives young people the skills they need to manage projects and communicate well, which sets them up to move up to the team leader and QA roles that are important to the internal training system.

## 3. Gender and Equity-Oriented Interventions

DDD's materials that are meant for the public always stress their commitment to inclusivity. According to DDD, about half of its employees around the world are women. The company aims to keep this balance in its Cambodian offices.<sup>28</sup> The organization also says that around 10% of its employees have disabilities. The focus on fairness is not an accident; it is a key part of the impact sourcing mission. Many news stories have talked about how DDD helps young women in Cambodia's tech sector get safe, formal jobs. This is an industry where young women have historically not been well represented. In 2013, for instance, a brief strike occurred over the implementation of a new performance-based pay structure, as reported by The Phnom Penh Post. The dispute centered on wages, with one 23-year-old female employee stating, "We work hard, but the salary we get is not enough to support our daily living and studies." The incident is notable not just for the dispute itself, which was resolved within days, but for the reported process: direct dialogue between management and staff representatives led to a mutually agreed-upon revision of the policy.<sup>29</sup> In this instance, while highlighting labor friction, it also demonstrates a mechanism for worker voice and responsive communication that is crucial to a functioning equity framework. The model's structure, which includes work, scholarships, and help with continuing education, may be especially beneficial for young women, who may feel more pressure from society or their families to skip college and get a job right away. Anecdotal stories often show how powerful this method can

be. In an article in Forbes, a former employee shared her dream to run an IT business; based on her experience at DDD, she made a plan for this.<sup>30</sup>

#### **4. Reported Socio-Economic Impacts**

DDD's impact reports emphasize notable improvements in participants' social and economic situations. On average, graduates of DDD increase their income by a factor of four within five years.<sup>31</sup> This number is very different from how wages usually go up in Cambodia. For example, in the garment/footwear/travel-goods (GFT) sector, minimum wage—nominal wages rose from \$190 (2020) to \$208 (2025), a +\$18 increase (~+9.5%) over five years.<sup>32</sup> According to the impact reports, DDD employees emphasize the value of educational opportunities through various DDD programs. The work-study model lets young people go to college while working at the same time.

DDD says that more than 2,000 young people in Cambodia have graduated from its program, and most of them attended college.<sup>33</sup> For many people, this mix of working and learning changes their lives. As one graduate mentioned in an interview, the combination of a job and a scholarship, enabled them to support themselves and their family."<sup>34</sup> The results that were reported are directly in line with a number of UN Sustainable Development Goals, especially SDG 1 (No Poverty), SDG 4 (Quality Education), SDG 5 (Gender Equality), and SDG 8 (Decent Work and Economic Growth).

## **V. Discussion and Conclusion**

In this single-case analysis from Cambodia, the train-the-trainer model appears to offer a promising pathway toward scalability as well as positive human development impact. DDD has built a self-sustaining ecosystem of skills by consistently investing in human capital. This

is very different from the extractive models of the gig economy. The results show that promoting all of middle management from within, offering extensive soft skills training, and staying committed to gender equity are all important parts of this success. The model not only provides the global AI supply chain with high-quality data, but it also has long-lasting and deep socio-economic benefits, such as helping Cambodian youth who do not have access to education and jobs to earn more money and do better in school.

The Digital Divide Data-inclusive "train-the-trainer" model thus compares quite favorably to crowdsourcing and BPO outsourcing. The micro-workers who use digital labor platforms for data-work have repeatedly been shown to suffer from poorly paid, abusive, often automated, managerial practices, such as arbitrary account deactivation, and wage theft from platform clients.<sup>35</sup> DDD's approach ensures that workers are treated with respect. In the BPO model, AI does not primarily destroy jobs, but redistributes them to poorly paid workers. In addition, despite their expertise, such data workers see their activity as 'menial' compared to highly paid and visible AI workers such as data scientists and software engineers.<sup>36</sup> The predominant Chinese model of inland sourcing through data labelling bases (DLBs), collaborative initiatives between tech companies and local organizations, is perhaps more fair to workers, although labor is still often paid poorly and there is relatively high turnover of college-educated staff.<sup>37</sup> DDD's model explicitly values its workforce.

However, a balanced evaluation of the DDD model requires acknowledging the limitations and future challenges facing this model. The world of AI is changing very quickly. The skills needed for data annotation are changing from basic tasks like labeling and categorizing to more difficult ones like those needed for Reinforcement Learning from Human Feedback (RLHF) and fine-tuning large

language models. This change is a constant challenge for the DDD model, which has to keep changing its training programs to meet the needs of the industry. As a result, the trainers themselves need to learn new skills as well. In the future, trainers will need to teach more than just technical skills. They will also need to teach logical reasoning, ethical judgment, and a high level of adaptability. This will require a constant investment in the highest level of pedagogical development.

This study is based on secondary sources and the authors' own experiences, which gives it a strong base. It also shows clear paths for future research. Since DDD has been around for more than 20 years, there is a great chance for long-term studies that are very rigorous. Such research could follow the long-term career paths of DDD graduates and give us numbers on how the model affects income, career advancement, and social mobility over a lifetime. Additionally, this chapter combines public information, but future research that includes in-depth qualitative interviews with current and former staff, trainers, and managers could provide more detailed and nuanced information. These studies could look at the real-life experiences of people who took part in the TtT model to find out about the subtle problems and successes that are not shown in official reports.

In conclusion, how is it possible for industry to recruit and train a skilled and dependable workforce that can provide the consistently accurate and nuanced data that advanced AI systems need? The main point of this chapter is that investing in people is not a side issue but a central, strategic need for building a responsible and strong AI future. The DDD model shows that it is possible to create a high-quality, scalable AI workforce while also promoting growth that includes everyone and gives people more power. Some outstanding research questions are how to scale this workforce cost-effectively to meet the needs of a very rapidly growing industry, as well as

how to bring a workforce up an ever steeper learning curve. As AI systems become more common in everyday life, the need for human-centered models that teach skills, leadership, and moral judgment will only grow stronger. The "train-the-trainer" method, which DDD is an example of, is one effective model for this important work.

## References

1. Nithya Sambasivan et al., "Everyone wants to do the model work, not the data work: Data Cascades in High-Stakes AI," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021).
2. Osunde A. Osoba and William Welser IV, *An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence* (Santa Monica, CA: RAND Corporation, 2017), [https://www.rand.org/pubs/research\\_reports/RR1744.html](https://www.rand.org/pubs/research_reports/RR1744.html).
3. Lilly Irani, "Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk," *South Atlantic Quarterly* 114, no. 1 (2015): 225–34.
4. Mary L. Gray and Siddharth Suri, *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass* (Boston: Houghton Mifflin Harcourt, 2019).
5. Victor S. Sheng, Foster Provost, and Panagiotis G. Ipeirotis, "Get Another Label? Improving Data Quality and Data Mining Using Multiple, Noisy Labelers," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2008).
6. M. J. Davda and D. V. Sindhu, "Impact of Human-in-the-Loop Paradigm in AI Labeling and Annotation," *Proceedings of the 2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)* (Bengaluru, India, 2024).
7. Jeff Howe, "The Rise of Crowdsourcing," *Wired Magazine*, June 14, 2006.
8. Irani, "Difference and Dependence among Digital Workers," 225–34.
9. Mary L. Gray and Siddharth Suri, *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass* (Boston: Houghton Mifflin Harcourt, 2019).
10. C. Le Ludec, M. Cornet, and A. A. Casilli, "The Problem with Annotation: Human Labor and Outsourcing between France and Madagascar," *Big Data & Society* 10, no. 2 (2023): –, <https://doi.org/10.1177/20539517231188723>.
11. Aneel Karnani, D. Miller, and D. Suri, "Impact Sourcing: A New Approach to Job Creation," *Ross School of Business Working Paper* No. 1180, 2012.
12. Jonathan Donner, Kentaro Toyama, and Katrin Verclas, "The Rise of Impact Sourcing," *Stanford Social Innovation Review*, 2015.
13. James J. Heckman and Tim Kautz, "Hard Evidence on Soft Skills," *Labor Economics* 19, no. 4 (2012): 451–64.
14. John Haskell, "The Train-the-Trainer Approach," *Journal of Extension* 40, no. 4 (2002).
15. Albert Bandura, *Social Learning Theory* (Englewood Cliffs, NJ: Prentice Hall, 1977).
16. Kristjan Blöndal and M. Ali Babar, "A Train-the-Trainer' Model for Scaling Software Engineering Education," in *Proceedings of the ACM/IEEE 44th International Conference on Software Engineering: Software Engineering in Practice* (2022).
17. National Institute of Statistics, Cambodia, *General Population Census of the Kingdom of Cambodia 2019* (Phnom Penh: National Institute of Statistics, 2019).
18. Royal Government of Cambodia, *Cambodia Digital Economy and Society Policy Framework 2021-2035* (Phnom Penh, 2021).
19. World Bank, *Cambodia Economic Update: Resilience and Reform* (Washington, DC: World Bank, 2021).
20. Robert K. Yin, *Case Study Research: Design and Methods*, 5th ed. (Thousand Oaks, CA: Sage, 2014).
21. Robert E. Stake, *The Art of Case Study*

Research (Thousand Oaks, CA: Sage, 1995).

22. Jenny Popay, Helen Roberts, and Amanda Sowden, "Guidance on the Conduct of Narrative Synthesis in Systematic Reviews," *ESRC Methods Programme*, 2006.
23. Global Impact Sourcing Coalition, "A Guide to Impact Sourcing," 2019.
24. Seline Chan, "How One Social Enterprise Has Disrupted The Cycle Of Poverty For 1,500 Young People," *Forbes*, October 27, 2016.
25. Digital Divide Data, "20 Years of Impact: Annual Report 2021."
26. Abt Associates, "What Works in Youth Skills Development? A Case Study of Digital Divide Data," 2014.
27. International Finance Corporation, *Digital Divide Data: A Case Study* (Washington, DC: IFC, 2012).
28. Digital Divide Data, "20 Years of Impact."
29. Mech Dara and Kevin Ponniah, "Digital Data Company Staff Strike Over Pay," *The Phnom Penh Post*, March 15, 2013
30. Chan, "How One Social Enterprise."
31. Abt Associates, "What Works in Youth Skills Development?"
32. KPMG Cambodia Ltd., *Technical Update*, October 2024, (Phnom Penh: KPMG Cambodia Ltd., October 2024), <https://kpmg.com/kh/en/home/insights/2024/10/technical-update.html>.
33. Digital Divide Data, "20 Years of Impact"; Abt Associates, "What Works in Youth Skills Development?"
34. Chan, "How One Social Enterprise."
35. Antonio A. Casilli, "Digital Labor Studies Go Global: Toward a Digital Decolonial Turn," *International Journal of Communication* 11 (2017): 3934–3954; Mary L. Gray and Siddharth Suri, *Ghost Work*.
36. C. Le Ludec, M. Cornet, and A. A. Casilli, "The Problem with Annotation."
37. Tongyu Wu, James Muldoon, and Bingqing Xia, "Global Data Empires: Analysing Artificial Intelligence Data Annotation in China and the USA," *Big Data & Society* 12, no. 2 (2025): -, <https://doi.org/10.1177/20539517251340600>.

## Author Positionality Statement

It is significant to communicate that the authors of this chapter currently work for Digital Divide Data Cambodia. Being "practitioner-researchers" gives us a different view of the case study. The authors' experience is not that of outside observers, but of people who have been deeply involved in the system this chapter is studying. The authors' jobs at DDD show a range of what the company does. The authors have worked in production management, team leadership, and learning and development, among other things. Several of us started out as entry-level operators and have personally moved through the career paths outlined in this chapter, including taking part in DDD's own "Training of Trainer" programs.

This study knows that this insider view could be seen as a possible source of bias. Instead, the authors relied upon extensive methodological approaches in organizational, self-ethnography, and action research design to contribute a reflexive and unique research perspective. This two-part research method of comprehensive secondary literary review combined with practitioner perspectives provides a multifaceted perspective on an organization, from its scope and mission to firsthand impact and challenges. The authors have a deeper, firsthand understanding of how things really work, how training is done, and how the workplace culture is than what can be learned from looking at public documents alone. So, this chapter combines a careful look at secondary sources with an interpretive lens based on years of direct involvement. The goal is to give a clear and unique view of the DDD model.



**Mr. Songheng Tieng**

*Songheng Tieng is in charge of training and development at Digital Divide Data (DDD) in Cambodia. He designs and runs the workforce development programs that are the main focus of this chapter. He is in charge of making sure that DDD's training programs include lessons on soft skills, computer literacy, and English, as well as coming up with the methods and tests that will be used to evaluate them. His career is an example of the "train-the-trainer" model because he has moved up from being a trainer to a leader. This path, along with his degrees in Social Sciences and certifications in HR and management, gives him a unique perspective on how the model works in the real world.*



**Mr. Sokheng Pok**

*Sokheng Pok is a Production Manager at Digital Divide Data (DDD). His career is a good example of the internal leadership pipeline that this chapter is about. He started as an entry-level Associate and worked his way up to Supervisor. Now he is in charge of daily operations, quality control, and team performance. He has led difficult AI data analysis projects for clients in the self-driving car industry. This hands-on look at DDD's career development path, which includes doing tasks, training others, and leading them, gives us a useful operational view of how the "train-the-trainer" model works in real life, as discussed in this research.*



**Ms. Chantola Khet**

*Chantola Khet has worked for Digital Divide Data (DDD) for more than ten years and is now in charge of Learning and Development. The "train-the-trainer" model is a great example of her career path. She began as a graphic editor and worked her way up to team leader, where she first taught her coworkers how to do technical things. Now, it's her job to make training materials and methods for computer and job skills. She finished DDD's own "Training of Trainer" program. She has a lot of experience in both production and teaching, so she has a well-informed and useful view on this research.*



# CASE STUDIES

**100**

Machine Translation System for Khmer Text to Khmer Braille  
*Kimhuoy Yann, Nab Mat, and Rottana Ly*

**110**

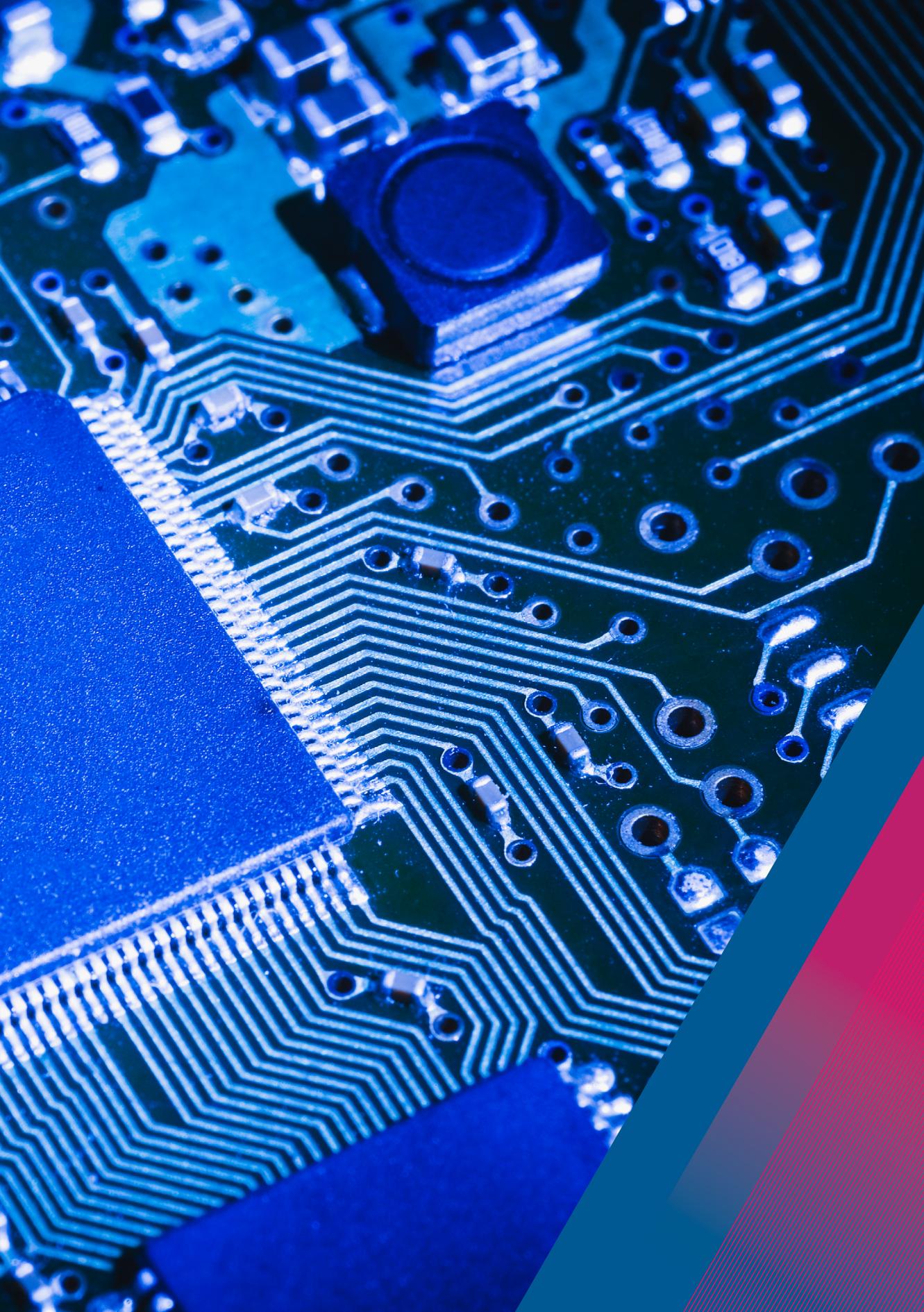
Reducing Antimicrobial Resistance (AMR) via Improving Microbiology Analysis Capacity and Optimizing the Use of Antibiotics through Machine Learning with Artificial Intelligence (AI)  
*Sophal Thear, Dr. Stanislas Rebaudet, Dr. Kennarey Seang, and Dr. Sotharith Bory*

**122**

HealthyRice: Enhancing Farmer-Expert Collaboration via Mobile Technology  
*Hongly Va and Soklong Him*

**134**

Policy-Driven Approaches in AI-Powered Verification for Combating Mis- and disinformation: A Cambodia-Focused Study  
*Prohim Tam and Riccardo Corrado*



# Machine Translation System for Khmer Text to Khmer Braille

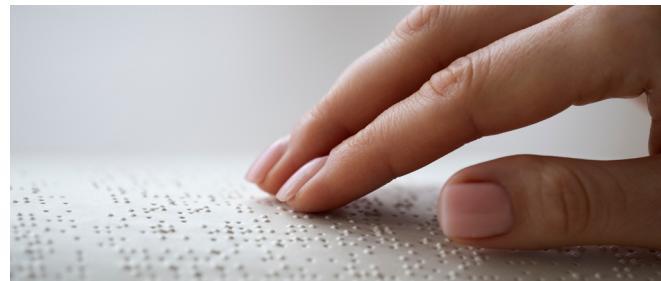


Kimhuoy Yann, Nab Mat, and Rottana Ly

## I. Introduction

People with visual impairments play a vital role in all communities. They are equally eager to explore and learn about the various aspects of their daily experiences.<sup>1</sup> Globally, an estimated 2.2 billion people live with some form of vision impairment, and at least 1 billion of these cases are preventable or untreated.<sup>2</sup> In Cambodia, around 144,000 people are blind, with about 28,800 new cases each year, mainly due to cataracts.<sup>3</sup> Meanwhile, many people in Cambodia who live with visual disabilities have limited access to assistive tools such as Khmer Braille learning materials and translation systems. This creates a strong need for solutions such as a Braille translation machine to support communication and education for visually impaired people in the country.

Although 80% to 90% of these cases can be prevented or treated, limited access to eye care and a shortage of trained professionals remain serious challenges, especially in rural areas, where 85% of the population lives. Given these conditions, there is a growing need for assistive tools, such as Khmer Braille translation systems, to support the education and communication needs of the visually impaired population. Although there is commercial Braille translation software, such as Duxbury Braille Translator,<sup>4</sup> it is expensive in Cambodia. As a result, only one printer at the NISE in Phnom Penh has a license, which limits access to many visually impaired users. This limited access to Braille translation tools significantly restricts educational and communication opportunities for visually impaired individuals in Cambodia. Consequently, there is a pressing



need for an affordable and locally adapted Braille translation system that can increase accessibility and support the inclusion of visually impaired people in education and daily life. To address this, we have developed a Khmer Braille translation machine designed to provide an affordable and efficient tool for converting Khmer text into Braille.

To develop a Braille translation system between Khmer text and Khmer Braille, we adopted a knowledge-based method that uses the Khmer Character Cluster (KCC) to accurately segment the text, followed by dictionary-based mapping and predefined conversion rules to generate Khmer Braille, which is then mapped to Braille Unicode.<sup>5</sup>

The remainder of this chapter is organized as follows. Section II outlines the key challenges related to visual impairment in Cambodia, while Section III introduces the structure and characteristics of Khmer Braille. Section IV describes the development of the Braille translation model, along with the outcomes and potential impact of the proposed system. Section V concludes the chapter, and Section VI discusses future work.

## II. Key Challenges

In this section, we discuss the main challenges faced by visually impaired individuals in Cambodia, particularly accessing information and participating in education. These challenges highlight the need for Khmer Braille translation tools and more inclusive systems to support equal opportunities for people with disabilities.<sup>6</sup>

### 1. Visual Impairment in Cambodia

Access to information is a fundamental right, yet in Cambodia, many people with disabilities, especially those with visual impairments, continue to face significant barriers. Although Cambodia has ratified the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD), practical challenges remain.<sup>7</sup> In addition, a recent report by the Cambodia Development Resource Institute (CDRI) highlights that access to essential information remains a critical issue, especially for people with visual impairments. Many rely on others to obtain news or opportunities, as accessible formats such as Khmer Braille, screen readers, or audio content are often unavailable. These limitations are more severe in rural areas, where assistive technology is scarce and digital literacy is low. According to the United Nations Development Program (UNDP) Cambodia, people with disabilities face multiple challenges, such as inequality and discrimination in access to education, healthcare, employment, social protection, justice, and public transport United Nations Development Program.<sup>8</sup> Furthermore, the NISE reported that only 10-20 students with disabilities continue to study higher education due to the lack of accessible materials and support, especially in Khmer Braille. This limits both their learning and long-term opportunities for social and economic inclusion.

### 2. Special Education and Inclusive School

Education is a basic human right and a vital path to lifting people out of poverty in a sustainable way.<sup>9</sup> It must be accessible to all children, including those with disabilities, to ensure equal opportunities for a better life. Despite efforts by the Cambodian government to promote inclusive education, significant challenges remain. These challenges include the lack of resources and funding, inadequate teacher training in inclusive practices, and social attitudes that perpetuate discrimination and marginalization.<sup>10</sup>

Since 2018, the Ministry of Education, Youth and Sport (MoEYS), with the support of UNICEF, has provided inclusive education training to teachers, benefiting more than 17,000 students with disabilities each year. A UNICEF-supported curriculum has also reached more than 500 marginalized learners. NISE, established in 2017, continues to build teacher capacity for inclusive classrooms.<sup>11</sup> Despite these efforts, access remains limited. In 2024, MoEYS reported that special education schools serve only 199 students with visual impairments, 184 with intellectual disabilities, and 635 with hearing impairments.<sup>12</sup> More than 57% of children with disabilities are still excluded from education, compared to just 7% of their non-disabled peers.<sup>13</sup> Barriers are especially pronounced in rural areas, where infrastructure and negative attitudes continue to restrict educational access.

### 3. Resources and Current Situation

According to information from NISE, Khmer Braille documents for people with visual impairments, especially students, can initially be produced using a slate and stylus with special Braille paper. Alternatively, we can use a Braille translator and a specialized printer, both of which are costly.

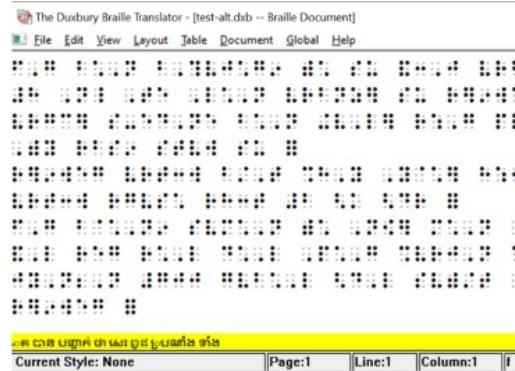
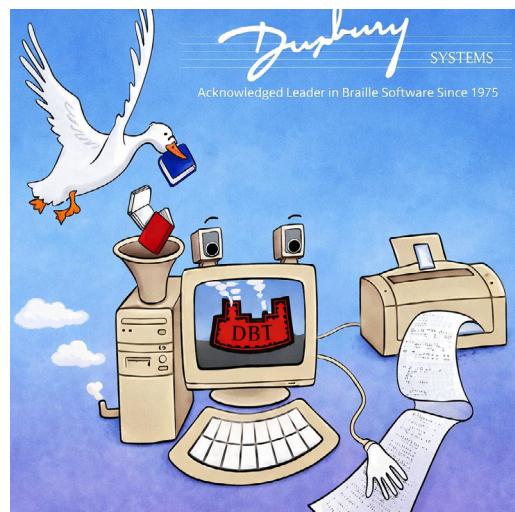
### 3.1 Slate and Stylus

The slate and stylus are the Braille writing tool (Figure 1); just like paper and pen, it is used by people with impairments so that they can use it to emboss dots on paper. Normally, it is used by the teacher or student who wants to learn Braille. As a challenge, it limits the number of Braille characters, making it time-consuming.



### 3.2 Braille Translator Software

Duxbury is one of the global Braille software translators that is used by most of the major Braille publishers and can translate more than 180 languages globally (Figure 2). However, when this program is used to translate text into Khmer Braille, we can identify several key challenges. Khmer complex script causes numerous issues with translation from Khmer text to Khmer Braille due to its complexity. The system also cannot render specific symbols in math, physics, chemistry, or biology, which requires a teacher who knows both Khmer Braille and those symbols or characters. With these problems, it becomes very time-consuming to produce a Braille document.



### 3.3 Braille Printer

After the translation from translator software, we need the printer to produce the Braille document, and as the current situation, we have only a few machines within the country since this specially designed printer is costly and is not affordable for all special education institutions. Figure 3.



Figure 3: Braille Printer

### III. What Is Khmer Braille?

#### 1. Background

According to interviews with teachers at NISE, Cambodian Braille was invented by Cambodians and Thais in 1988 in the Saitou camp and uses raised dots (⠼) to represent characters.<sup>14</sup> Derived from Khmer script, Cambodian Braille is based on a 6-dot character system, arranged in a grid of two parallel vertical lines of three dots each, which allows for 128 possible characters. These 6-dot cells are read top to bottom and left to right, enabling users to feel and interpret written information, much like reading a sighted book. In 1994, the Khmer Braille system was further refined by the Krousar Thmey community, in collaboration with the MoEYS and the Maryknoll Lay Missioners, to better align with the structure of written Khmer.<sup>15</sup> Given the complexity of the Khmer writing system, which contains more than 144 characters, the standard 6-dot Braille system, with its 63-character limit, proved insufficient. To address this, Cambodian Braille innovatively employs an 18-dot system, formed by combining three 6-dot cells to represent a single character, thereby extending its capacity to encompass the full range of Khmer characters. Two groups have been separated in the Khmer consonants based on their sound. The A series has the sound of aw, and the O series has the sound of oh. In the Braille pattern, there is a general rule for most of the O-series Braille consonants coming from the A-series. By adding this cell (.) to the A series, we can form the O series, and there are some special cases of unique O series that do not come from the A series, and to form a sub-consonant is just add (v) before its consonant. The examples are shown in Figure 4 below.

A-series	O-series
ក ⠼	គ ⠼⠼
ខ ⠼	យ ⠼⠼⠼
ច ⠼	ជ ⠼⠼⠼
Special case	
ឃ	⠼
ុ	⠼
ុ	⠼
ុ	⠼
ុ	⠼
ុ	⠼⠼
ុ	⠼⠼⠼
Consonant	
ក ⠼	ន ⠼⠼
ខ ⠼	ន ⠼⠼⠼
ច ⠼⠼	ន ⠼⠼⠼⠼

Figure 4. O-series Braille Consonants

#### 2. Structure

In addition to consonants, there are vowels, independent vowels, punctuation, and some special signs used in the Khmer language, and each category has a unique combination of Braille cells from one to three cells representing one element, with no general term, unlike consonants. In written form, vowels are represented as diacritics, whereas in Braille, they are considered complete letters and adhere to the position of the corresponding vowel in the printed text. For instance, in the word ឃុំ (Khmer) in printed text. The vowel ុ (e) comes before the consonant cluster ឃ (khm). However, in Braille, the vowel (e) follows the consonant cluster (khm), resulting in the representation (ekhmr). Khmer Braille represents each Khmer character using one to three Braille cells, depending on the symbol. Table 1 lists all Khmer characters, and Figure 4 shows their Braille equivalents.

Table 1: Summary of Khmer Characters

## IV. Project Development

## 1. Knowledge-Based System

With the complexity of the Khmer language, there are rules and special cases during the translation that cause errors. To build a knowledge-based machine translation system for Khmer Braille, we have developed a dictionary mapping between Khmer individual characters and Braille ASCII printable. We will present the special key with this dictionary.

A. Dictionary Mapping: We have constructed a dictionary that maps Khmer characters to ASCII printable characters. This includes consonants, sub-consonants, vowels, independent vowels, punctuation marks, and additional characters used in Khmer text.

- B. Cluster Segmentation: We need to perform a special segmentation called KCC segmentation at the word level to accurately capture each word in Khmer.
- C. Order Handling: In the KCC format, some vowels (like ឃ, ង, ឃោ, ឃុោ) appear before the consonant in print but must be reordered in Braille.
- D. Vowel Combinations: There are also special combinations of vowels that can cause confusion during the mapping process with the dictionary. Examples include: ោ + ែ = ោែ, ោ + ែោ = ោែោ, ឃ + ែោ = ឃោែោ, ឃ + ោោ = ឃោោោ. So, we need to change its order.
- E. Number Cases: In Braille, there is a special case with the number sign, which needs to be placed before the number. The challenge arises when the number has more than two digits, as only one sign is required. Another special case is in float numbers: the number sign will appear only for the integer digit, meaning there is no number sign for the fractional part. This needs to be carefully checked and confirmed during the mapping process.

For example, the Khmer word {ឃុំ} is composed as {ឃ + ុ + ុ + ុ}. However, in Khmer Braille (ASCII representation), the correct reading order is {ុ + ឃ + ុ + ុ}. This reflects the structural rule in Braille that vowel signs appearing before the consonant are written first.

## 2. Web-Based Development

Our system is a web application-based system that we already deploy as production for the users, especially teachers from NISE. Below is the detailed process that we developed for our model system.

- Frontend Developed using React.js to deliver a responsive, user-friendly interface that supports dynamic interactions.
- The backend is powered by Flask, providing lightweight and scalable REST APIs for processing translation tasks.
- The entire application is containerized using Docker and integrated with GitHub for version control and CI/CD automation, ensuring easy updates and reproducibility across environments.
- Outcome and Impact** After building the knowledge-based machine translation system for Khmer Braille and developing this system into web-based applications, we are able to launch this system for public use. Currently, the system has been tested by teachers from the NISE, and we can achieve up to 99% accuracy for the translation from Khmer text to Khmer Braille and vice-versa, in general context. Below are the functionality and guidelines of our Khmer Braille Translator System, Figure 5.
  - The Home Page is the page that takes us to the homepage of the website.
  - Feedback is the page where users can submit their comments to our website.
  - Help is a page that shows how to use the website for translating into Braille.
  - About Us is the page that shows the origin of the website and its creators.
  - Translate Khmer Document to Braille is a button that we use to upload documents for translation.
  - Math Symbol: We have developed some math symbols but have not completed all of them yet.

*Figure 5: Interface of Khmer Braille Machine Translation System*

Figure 6: The Output of Khmer Braille Machine Translation System

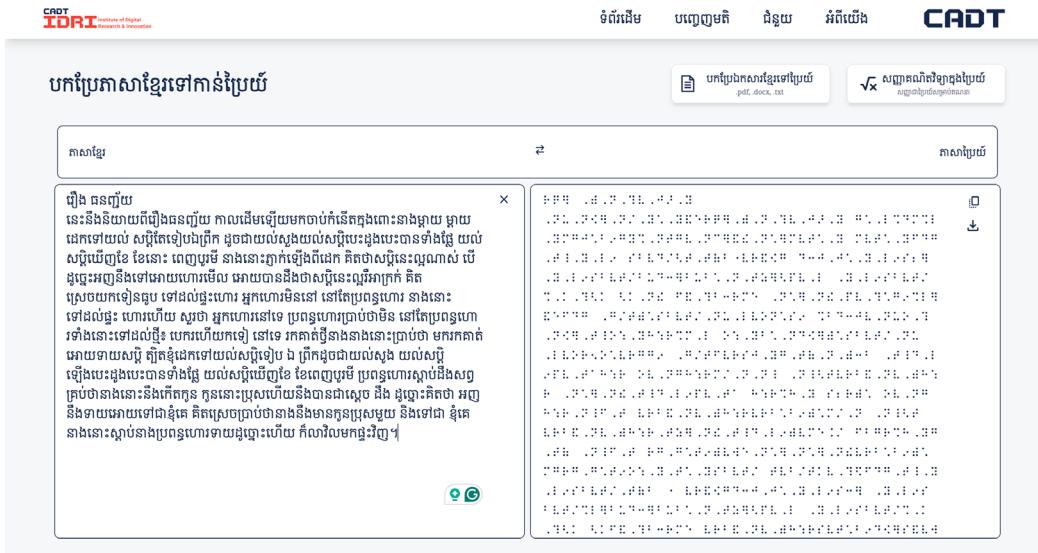


Figure 6 shows an example output of the Khmer Braille machine translation system. Users can either type Khmer text directly using the keyboard or paste or upload it into the text box, and the system will automatically convert the Khmer text into Khmer Braille. After the translation, users can copy the Braille text or download it as a text file or Word document.

## V. Conclusion

In this chapter, we have built a knowledge-based machine translation system to translate Khmer text into Khmer Braille. The system showed strong performance by effectively capturing the linguistic structure of Khmer Braille, which follows strict grammatical rules. Its clear and straightforward design makes it well-suited for accurately representing the complexity of the Khmer language. These results highlight the potential of this approach to improve access to written information and educational materials for the visually impaired community in Cambodia. By providing a reliable and understandable way to convert Khmer text into Braille, this approach can contribute to greater inclusion and accessibility in education and everyday life.

## VI. Future Work

Currently, the system has been deployed and used by the teachers from NISE, and we plan to launch this project to other special education high schools, particularly schools in provinces such as Battambang, Siem Reap, and Kampong Cham. In the meantime, we plan to connect this system with Braille printers to make it easier for teachers from special education schools to produce educational materials. This would help speed up the creation of customized Braille documents and reduce the workload involved in preparing resources for visually impaired students. We also want to integrate the system with Braille display devices, so users can access content digitally without needing to print it. This would be especially helpful for families who have the devices and prefer a more convenient option. On top of that, we plan to add speech-to-text capabilities and expand the system to support multiple languages, making it more useful and accessible to a wider audience. The proposed system is designed to address the communication gap experienced by the visually impaired community. By using digital technologies, we aim to create meaningful

social impact, thereby promoting greater equity and inclusion across various domains in our society.

## **VII. Acknowledgements**

This chapter was initiated by Dr. Ye Kyaw Thu and was made possible through the collaboration between NISE and CADT. We would like to thank teachers from NISE for providing essential resources, Khmer Braille instruction and rules, as well as for manually verifying the dataset. We also extend our gratitude to CADT researchers, including Mr. Veng Ponleur and Mr. Samol Raingsey, for their contribution to this chapter.

## References

- Shokat, Sana, Rabia Riaz, Sanam S. Rizvi, Khalil Khan, Farina Riaz, and Se J. Kwon. 2022. "Analysis and Evaluation of Braille to Text Conversion Methods." *Wiley Online Library*. <https://onlinelibrary.wiley.com/doi/10.1155/2020/3461651>.
- Hill, Alyssa. 2025. "Blindness Statistics 2025." *Vision Center*. <https://www.visioncenter.org/resources/blindness-facts/>.
- Sogbesan, Enitan, and Uch Yutho. n.d. "Cambodia's National Eye Care Programme and VISION 2020: The Right to Sight." *National Library of Medication*. Accessed 2025. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1705982/>.
- "About Duxbury Systems." n.d. *Duxbury Systems*. Accessed October 16, 2025. <https://www.duxburysystems.com/about.asp>.
- Buoy, Rina, Nguonly Taing, and Sokchea Kor. 2020. "Khmer Word Segmentation Using BiLSTM Networks." (December). [https://www.researchgate.net/publication/347516369\\_Khmer\\_Word\\_Segmentation\\_Using\\_BiLSTM\\_Networks](https://www.researchgate.net/publication/347516369_Khmer_Word_Segmentation_Using_BiLSTM_Networks).
- Veng, Ponleur, Kimhuoy Yann, and Rottana Ly. 2024. *In Proceedings of the 6th ACM International Conference on Multimedia in Asia Workshops*, 1-5. N.p.: Association for Computing Machinery. <https://doi.org/10.1145/3688578.3688903>.
- Chea, Phal, and Phoury Bun. 2024. "Access to Information for Persons with Disabilities in Cambodia: Challenges and the Way Forward." *CDRI*. <https://cdri.org.kh/news/access-to-information-for-persons-with-disabilities-in-cambodia-challenges-and-the-way-forward>.
- United Nations Development Programme (UNDP) Cambodia. n.d. "Programme to Promote Disability Inclusion (PPDI)." *United Nations Development Programme*. Accessed 2025. <https://www.undp.org/cambodia/projects/programme-promote-disability-inclusion-ppdi>.
- Cambodian Education Forum. 2023. "Inclusive Education for Children with Disabilities in Cambodia: Progress, Challenges, and Ways Forward." (09). <https://cefcambodia.com/2023/09/18/inclusive-education-for-children-with-disabilities-in-cambodia-progress-challenges-and-ways-forward/>.
- Lem, Chantou. 2021. "Towards Inclusive Education in Cambodia: Overcoming Barriers and Embracing Diversity." *Journal of Education Review Provision (JERP)* 1, no. 3 (12): 101-109. [10.55885/jerp.v1i3.205](https://doi.org/10.55885/jerp.v1i3.205).
- "Inclusive Education in Cambodia: Successes and Challenges." 2024. *Unicef*. <https://www.unicef.org/cambodia/stories/inclusive-education-cambodia-successes-and-challenge>.
- ACCESS 2 (Australia-Cambodia Cooperation for Equitable Sustainable Services). 2024. "Enhancing Disability Inclusion, Equity and Rights in Cambodia: A Resource for Action." *ACCESS 2*. <https://access2cambodia.org/access-2-publications/>.
- UNESCO Institute for Statistics (UIS). 2017. "Education and Disability." *UNESCO*. <https://uis.unesco.org/en/news/children-disabilities-are-more-likely-be-out-school>.
- Veng, Ponleur, Kimhuoy Yann, and Rottana Ly. 2024. *In Proceedings of the 6th ACM International Conference on Multimedia in Asia Workshops*, 1-5. N.p.: Association for Computing Machinery. <https://doi.org/10.1145/3688578.3688903>.
- Ministry of Education, Youth and Sport (MoEYS), Cambodia. 2020. "Khmer Braille." <https://moeys.gov.kh/>.



**Ms. Kimhuoy Yann**

*Kimhuoy YANN completed her Computer Science B.E degree at Institute of Technology of Cambodia (ITC) in 2023 and is now pursuing a master's degree in computer science at CADT. She is also employed as a lecturer-researcher assistant at CADT. She is focused on studying natural language processing and computer vision*



**Ms. Nab Mat**

*Nab Mat received her B.S. degree in Computer Science and Engineering from the Royal University of Phnom Penh, Cambodia, in 2021, and her M.S. degree in ICT Convergence from Soonchunhyang University, South Korea, in 2024. Her research interests include human activity recognition, computer vision, image processing, machine learning, and deep learning applications. She is currently a full-time lecturer and researcher at CADT.*



**Dr. Rottana Ly**

*Dr. Rottana Ly is currently Director of Research and Development at CADT. He has been a lecturer and researcher since 2017. Rottana brings a strong academic and research background to his current role. He holds a Ph.D. of Artificial Intelligence, with a specialization in Computer Vision, from the University of Grenoble, France, and a Master's degree of Computer Science from University of Paul Sabatier, France. His research interests focus on applications of AI and Machine Learning in different domains including education, healthcare, and agriculture.*

# Reducing Antimicrobial Resistance via Improving Microbiology Analysis Capacity and Optimizing the Use of Antibiotics through Machine Learning with Artificial Intelligence

---



*Sophal Thear, Dr. Stanislas Rebaudet, Dr. Kennarey Seang, and Dr. Sotharith Bory*

## I. Introduction to Antimicrobial Resistance and Its Global Impact

Antimicrobial Resistance (AMR) is a growing global threat where common infections are becoming untreatable. It represents a serious and escalating threat to global public health, significantly undermining the effectiveness of modern medicine. It occurs when microorganisms such as bacteria, viruses, fungi, and parasites evolve to resist the effects of antimicrobial drugs (e.g., antibiotics, antivirals, antifungals), making infections harder or impossible to treat.<sup>1</sup> This natural evolutionary process is badly accelerated by the widespread and often inappropriate use of antimicrobials in human medicine, animal agriculture, and the environment, creating selective pressure that favors resistant strains.<sup>2</sup>

Socially, AMR exacerbates health inequalities, extremely affecting low- and middle-income countries where access to effective antimicrobials, diagnostic tools, and infection prevention and control measures may be limited.<sup>3</sup> It threatens food security through its impact on animal health and agriculture, and can push vulnerable populations into extreme poverty due to high treatment costs and loss of income. The complex interplay of social, economic, and environmental factors highlights the need for a comprehensive “One Health” approach to combat AMR, recognizing

that the health of humans, animals, and the environment are interconnected.<sup>4</sup> Economically, AMR raises a significant burden on healthcare systems and national economies. Treating resistant infections often requires more expensive, last-resort drugs, longer hospital stays, and more intensive care, driving up healthcare costs.<sup>5</sup>

Infections that were once easily curable can become untreatable, resulting in severe illness, disability, and death. For instance, in 2019, an estimated 4.95 million deaths worldwide were associated with bacterial AMR, with 1.27 million directly attributable to it.<sup>6</sup> Projections indicate that the direct healthcare costs associated with AMR could rise significantly in the coming decades, potentially costing hundreds of billions of dollars annually.<sup>7</sup> Beyond healthcare expenditures, AMR contributes to productivity losses due to prolonged illness, disability, and early deaths, affecting workforces and economic growth globally.

## II. Clinical Decision Support Systems in Infectious Disease Management

Clinical Decision Support Systems (CDSS) have emerged as powerful tools in infectious disease management, particularly in addressing the complexities of diagnosis, treatment, and prevention. These systems leverage clinical guidelines, patient data, and emerging

evidence to provide healthcare providers with actionable recommendations. The integration of CDSS into clinical workflows has been shown to improve adherence to guidelines, optimize antibiotic use, and enhance patient outcomes.

These systems are particularly valuable in scenarios where clinical decisions must be made rapidly, such as during outbreaks or when managing antibiotic resistance. One of the most significant contributions of CDSS is in antibiotic stewardship. Studies have shown that CDSS can reduce inappropriate antibiotic use by guiding clinicians toward the most effective and least harmful treatment options.<sup>8</sup> For instance, a systematic review of CDSS implementations found that these systems consistently improved guideline adherence and reduced antibiotic consumption.<sup>9</sup> This is particularly important in the context of rising antibiotic resistance, where optimizing antibiotic use is critical to preserving their effectiveness.

Outside regular care, CDSS have also proven invaluable in the management of emerging infectious diseases. For example, during the COVID-19 pandemic, CDSS were used to support diagnosis, treatment, and monitoring. These systems helped clinicians interpret complex guidelines and adapt to rapidly evolving evidence.<sup>10</sup> Similarly, a CDSS tool was developed specifically for the evaluation of suspected monkeypox infections, enabling clinicians to triage patients effectively and ensure timely interventions.<sup>11</sup>

An AI-CDSS developed for detecting ceftazidime-avibactam resistance in *Klebsiella pneumoniae* exemplifies the potential of AI in infectious disease management.<sup>12</sup> This system combined matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) with machine learning algorithms to rapidly identify resistance patterns. Such systems can accelerate clinical decision-making and improve treatment outcomes.

### III. Machine Learning Applications in Predicting AMR

The integration of artificial intelligence into healthcare constitutes a significant step forward, ushering in a new era of precision medicine.<sup>13</sup> The ability of AI to handle massive volumes of data and translate it into meaningful knowledge and insight pushes data analysis well beyond human cognitive boundaries. This capability addresses challenges associated with insufficient information and time constraints that often hinder rational decision-making. Machine learning algorithms, by design, learn objectively, frequently yielding predictions more accurate than those made through traditional clinical practice.<sup>14</sup> Various machine learning models have been applied to the prediction and classification of AMR.<sup>15</sup>

Machine learning (ML) algorithms offer significant potential in combating AMR by enabling accurate prediction of resistance patterns and facilitating targeted antibiotic use. Researchers have successfully applied diverse ML models, including decision trees, XGBoost, and Random Forest classifiers.<sup>16</sup> Beyond traditional models such as Logistic Regression, which served as a foundational baseline, recent advancements in supervised machine learning have significantly enhanced predictive capabilities in susceptibility analysis. Notably, powerful ensemble techniques such as Extreme Gradient Boosting (XGB) and Random Forest (RF) have demonstrated superior performance. XGB, a highly regularized form of gradient boosting, excels at controlling overfitting while iteratively building strong predictive models,<sup>17</sup> whereas RF improves upon bagging by decorrelating individual decision trees, leading to more robust and generalized predictions.<sup>18</sup> Furthermore, Artificial Neural Networks (NNs) represent a cutting-edge approach, employing a dense multilayer architecture with input, multiple hidden, and output layers. These networks leverage complex activation functions like ReLU and

incorporate regularization techniques such as batch normalization and dropout layers to capture intricate feature relationships and mitigate overfitting, positioning them as highly capable models for complex prediction tasks.<sup>19</sup>

#### IV. Data Privacy and Ethical Considerations in Healthcare AI

Healthcare AI systems typically rely on large volumes of patient data, including electronic health records (EHRs), lab results, and medical imaging. Such data is subject to strict privacy regulations due to its sensitivity. One of the key privacy concerns lies in how AI models require extensive data to perform accurately, increasing the risk of breaches or misuse. Even when anonymization techniques are applied, re-identification remains a threat, and patients often lack clarity on how their data are used in model training or decision-making processes. Weak data governance and the lack of clear consent practices further increase these risks.<sup>20</sup>

In addition to privacy, healthcare AI presents ethical concerns related to fairness, transparency, accountability, and patient autonomy. A key issue is unintended algorithmic bias. When AI models are trained on datasets that are unbalanced or not representative of the whole population, the results can be unfair. This can lead to inaccurate diagnoses or unequal treatment across patient groups. Bias in AI can deepen existing health disparities, especially among vulnerable communities. To reduce these risks, developers must use diverse datasets, regularly evaluate models for fairness, and apply proper validation methods.<sup>21</sup>

Transparency and explainability are also critical in clinical settings, where “black box” AI models may provide accurate predictions but fail to offer interpretable reasoning. When neither doctors nor patients can understand an AI’s reasoning, trust in the system declines. This lack of explainability makes it difficult to use AI in real-world healthcare, especially in serious cases like disease diagnosis or treatment decisions, where clarity is essential.<sup>22</sup>

Accountability is another important concern. When an AI system makes an incorrect diagnosis or suggests an ineffective treatment, determining responsibility becomes complex. There is currently no universally accepted legal framework for assigning liability in such cases, which creates ethical ambiguity. Without clear roles and responsibilities, there is a risk that no one will be held accountable for patient harm. Ensuring clear chains of responsibility between developers, institutions, and practitioners is necessary for ethical deployment.<sup>23</sup>

#### V. Case Study: AI-Driven Application of AMR Management in Cambodia

The project aims to leverage ML techniques to analyze AMR data and develop an AMR predictive model from locally input data based on machine learning techniques. This initiative will aid doctors and clinicians in understanding resistance patterns and making informed treatment decisions.

This case study details the development of a web application designed to predict antibiotic susceptibility, a crucial tool in combating AMR. The project involved a systematic approach from data acquisition and preparation to model selection and application development, with a strong emphasis on ethical considerations.

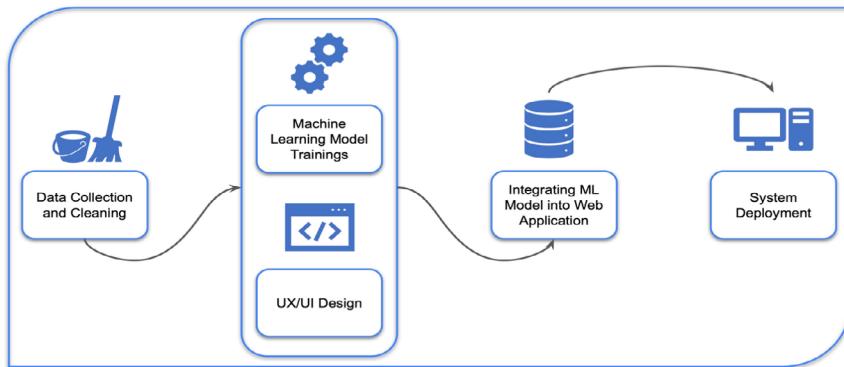


Figure 1: Flowchart of the CAMPRA Application

## 1. Data Collection and Ethical Considerations

The initial and foundational phase of this project involved extensive data collection from Calmette Hospital, conducted in collaboration with the University of Health Sciences and Cambodia Academy of Digital Technology. This critical dataset comprised two main components: *antibiogram data*, which provides information on the susceptibility of various bacterial isolates to different antibiotics, and relevant patient *demographic data*.

Following ethical guidelines was a crucial part of this data collection process. All personally identifying patient information, including names in both Khmer and English, was carefully eliminated from the dataset in order to protect patient privacy and adhere to data protection guidelines. This anonymization process was a non-negotiable step, protecting sensitive patient data throughout the project lifecycle.

## 2. Data Cleaning and Transformation

Upon collection, the raw data underwent a thorough process of cleaning and transformation. This step was essential to prepare the dataset for machine learning model training, addressing issues such as missing values, inconsistencies, and formatting errors. Data cleaning involved identifying and correcting errors, while transformation

focused on structuring the data into a format suitable for algorithmic processing. This cautious preparation ensured the integrity and reliability of the data, which are vital for accurate model predictions.

## 3. Machine Learning Model Development and Evaluation

With the cleaned and transformed dataset, the next phase involved applying various ML and deep learning (DL) models to predict antibiotic susceptibility. A diverse range of algorithms was tested to identify the most effective predictive model for the specific context of AMR data. The models explored included:

### Traditional Machine Learning Models

- Random Forest: An ensemble learning method that constructs a multitude of decision trees.
- Logistic Regression: A statistical model used for binary classification.
- Gradient Boosting: A powerful ensemble technique that builds models sequentially.
- XGBoost: An optimized distributed gradient boosting library designed for efficiency, flexibility, and portability.

## Deep Learning Models

- Bayesian Neural Network: A type of neural network that incorporates Bayesian inference to quantify uncertainty in predictions.
- Multi-Layer Perceptron (MLP): A class of feedforward artificial neural network.

Each model's performance was thoroughly evaluated using a comprehensive suite of metrics to provide an understanding of their predictive capabilities. These evaluation metrics included:

- Overall Accuracy: The proportion of correctly classified instances.
- Label-Based Accuracy: Accuracy calculated for each specific antibiotic susceptibility label (e.g., susceptible, intermediate, resistant).
- Area Under the Receiver Operating Characteristic (AUC ROC) Curve: A measure of a model's ability to distinguish between classes.
- Precision and Recall Curve: Illustrates the trade-off between precision (proportion of true positive predictions) and recall (proportion of actual positives correctly identified) for different thresholds.

Model	Metric	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Random Forest	Overall Accuracy	0.41	0.43	0.43	0.45	0.45
	Label-based Accuracy	0.85	0.87	0.88	0.90	0.90
Logistic Regression	Overall Accuracy	0.21	0.22	0.25	0.24	0.24
	Label-based Accuracy	0.76	0.78	0.82	0.82	0.83
Gradient Boosting	Overall Accuracy	0.36	0.37	0.37	0.36	0.36
	Label-based Accuracy	0.86	0.86	0.86	0.86	0.86
Extreme Gradient Boosting	Overall Accuracy	0.38	0.38	0.37	0.37	0.38
	Label-based Accuracy	0.86	0.86	0.86	0.86	0.86
Bayesian Neural Network	Overall Accuracy	0.80	0.83	0.84	0.85	0.84
	Label-based Accuracy	0.80	0.83	0.84	0.85	0.84
Multi Layer Perceptron	Overall Accuracy	0.24	0.26	0.31	0.31	0.33
	Label-based Accuracy	0.80	0.83	0.84	0.84	0.85

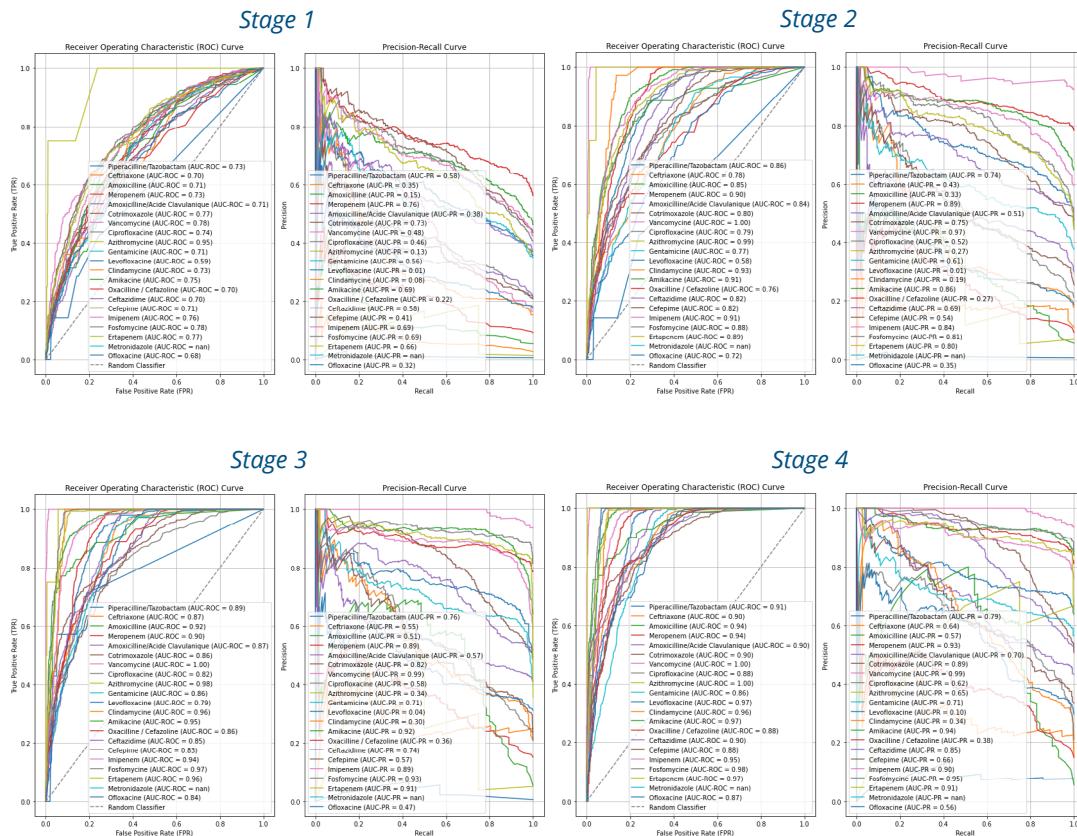
The table compares the performance of six machine learning models across five stages using two metrics: *Overall Accuracy* and *Label-based Accuracy*. The Bayesian Neural Network achieved the highest Overall Accuracy, with values ranging from 0.80 (Stage 1) to 0.85 (Stage 4). Logistic Regression consistently reported the lowest Overall Accuracy, ranging from 0.21 to 0.25. For Label-based Accuracy, the Random Forest model showed the highest performance, reaching 0.90

in Stages 4 and 5. The other models (Gradient Boosting, Extreme Gradient Boosting, Bayesian Neural Network, and Multi Layer Perceptron) showed high Label-based Accuracy values, generally between 0.80 and 0.86.

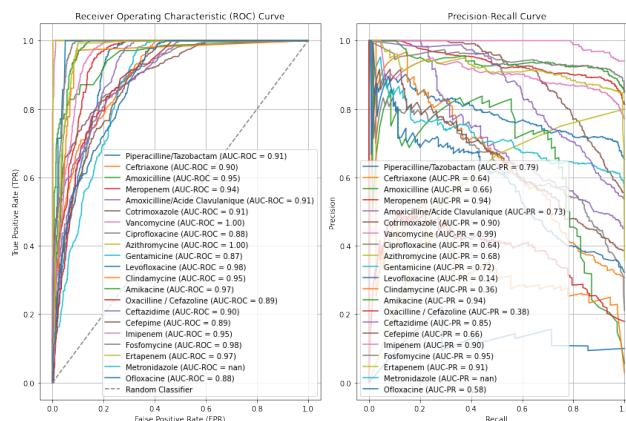
#### 4. Model Selection and Web Application Integration

Following the extensive evaluation phase, the *Random Forest* model was selected as the optimal choice for embedding into the web application. This decision was primarily driven by its consistently strong performance across all key evaluation metrics (Overall Accuracy, Label-Based Accuracy, AUC ROC, and Precision and Recall curves). The Random Forest model demonstrated a robust ability to predict antibiotic susceptibility, making it a reliable foundation for clinical decision support.

Model	Metric	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
<b>Random Forest</b>	Overall Accuracy	0.41	0.43	0.43	0.45	0.45
	Label-based Accuracy	0.85	0.87	0.88	0.90	0.90



## Stage 5



The figures show the Random Forest model's performance over five training stages for predicting antibiotic sensitivity, evaluated using Receiver Operating Characteristic (ROC) curves and Precision-Recall (PR) curves. Performance is quantified by the Area Under the Curve (AUC-ROC and AUC-PR) for each antibiotic. A performance improvement is observed across the stages. For AUC-ROC, scores generally increase; for instance, Piperacillin/Tazobactam rises from 0.73 (Stage 1) to 0.91 (Stage 5), and several antibiotics achieve 1.00 like Vancomycin. A similar upward trend occurs in AUC-PR, with Piperacillin/Tazobactam improving from 0.58 (Stage 1) to 0.79 (Stage 5). The consistent increase in both AUC metrics indicates that the Random Forest model's ability to discriminate between outcomes is enhanced with successive training stages.

The selected Random Forest model was then integrated into the web application. This application is designed to serve as an intuitive and accessible bedside tool for clinicians. By inputting relevant patient and infection data, the application leverages the embedded ML model to provide rapid insights into antibiotic resistance and sensitivity probabilities, thereby guiding more informed and targeted antibiotic prescription decisions in the fight against AMR.

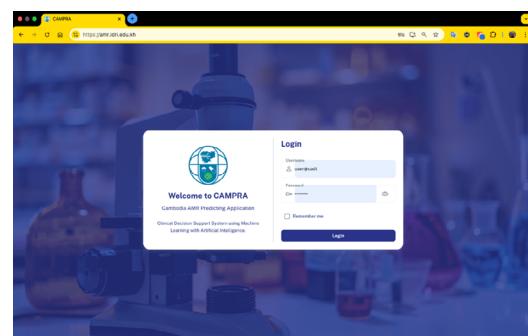


Figure 2: CAMPRA Sign in Page

Figure 3: CAMPRA Form Inputs

Figure 4: CAMPRA Inputs Filled with Data input

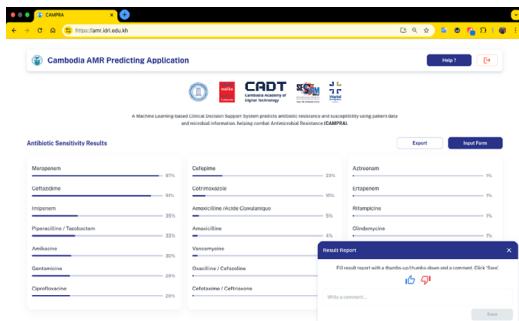


Figure 5: CAMPRA Prediction Results

## 5. How Application Work

The procedure requires completion of the Patient Form and the Stage Form. The Patient Form requires seven mandatory fields: Ward, Gender, Age, Address, Date, Sample, Diagnosis and Infection Type. The Stage Form defines a sequential process with field dependencies: Direct Observation must be selected first, which enables Culture; Culture must be selected to enable Genre; and Genre must be selected to enable Species.

After completing the forms, data submission is initiated. The system validates all required fields and dependencies, preventing submission if inputs are invalid or missing. Successful submission triggers the prediction model, which generates and displays antibiotic sensitivity results as percentages (probability). If the prediction probability is above 50%, the result is classified as *Sensible*, and if it is under 50%, it is classified as *Resistant*. Doctors and clinicians must independently analyze and validate this result, using their own clinical experience and knowledge to make a final decision for patient prescription and treatment.

Additionally, this application is designed exclusively for use by qualified doctors, clinicians, and healthcare professionals. It is a professional clinical tool, and its access and use are restricted to licensed medical personnel. It is not intended for the general public, patient use, or self-diagnosis, as its outputs require interpretation by a medical professional.

## VI. Opportunities and Challenges

Developing a web application for predicting antibiotic susceptibility presents both significant opportunities and inherent challenges.

There are many opportunities this tool creates. It helps doctors make better, faster decisions about which antibiotic to prescribe by giving them quick, data-driven advice. This supports careful antibiotic use, guiding doctors to the best choice and helping slow down the spread of drug resistance. The app can also use a patient's specific data and local resistance information to recommend treatment tailored just for them, which helps patients get better results with fewer side effects. It also makes hospital work more efficient by reducing the need to wait for slow lab tests, which could even help patients go home sooner. Finally, because it's a web-based platform, it is easy for doctors to access and can be easily expanded to new regions or hospitals.

However, there are also challenges. First, the model needs good, complete data to work well, and this data can be hard to get. Also, a model trained at one hospital might not work well at another because patients and local bacteria are different. We must also protect patient privacy and keep their data secure, which is a constant concern. Plus, bacteria and resistance patterns always change, so the model must be updated all the time with new data to stay accurate. Finally, doctors need to trust the advice, so a key challenge is making it easy to understand why the model recommends a specific antibiotic.

## VII. Strategic Recommendations

To maximize the impact and overcome the challenges associated with the web application for predicting antibiotic susceptibility, several strategic recommendations are crucial for future development and implementation.

These include: establishing robust data governance and secure sharing protocols to enhance data quality and generalizability; implementing continuous model monitoring and retraining with new local data to maintain accuracy; fostering interdisciplinary collaboration among data scientists, clinicians, and ethicists to ensure practical utility and address ethical concerns; integrating Explainable AI (XAI) techniques to provide transparent model insights and build clinician trust; exploring advanced privacy-preserving AI methods like federated learning to overcome privacy concerns and improve model generalizability across institutions; and working towards seamless integration with existing EHR and laboratory information systems (LIS) to streamline workflows and minimize manual data entry.

## VIII. Conclusion

The development of this web application for predicting antibiotic susceptibility represents a significant step forward in leveraging machine learning to combat antimicrobial resistance and in the context of applying AI in the healthcare sector in Cambodia. By providing clinicians with a data-driven tool for informed antibiotic prescription, the project aims to enhance treatment efficacy, promote responsible antibiotic stewardship, and ultimately improve patient outcomes. While the journey involves navigating complex challenges related to data, ethics, and clinical

integration, the strategic recommendations outlined provide a roadmap for continuous improvement and broader impact. This initiative underscores the transformative potential of interdisciplinary collaboration and technological innovation in addressing one of the most pressing global health threats.

## References

1. World Health Organization. (2023). *Antimicrobial resistance*.
2. Tizhe, D. T., Ogra, I. O., Apollos, S. D., Enesi, K. O., Yohanna, A., Aminu, R., & Kwaga, J. K. P. K. (2024). *Antimicrobial resistance as a global public health threat: the way forward*. *EUREKA: Life Sciences*, (1), 1-13.
3. *Ibid.*
4. *Ibid.*
5. O'Neill, J. (2016). *Tackling drug-resistant infections globally: final report and recommendations. The Review on Antimicrobial Resistance*.
6. Murray, C. J. L., Ikuta, K. S., Sharara, F., Swetschinski, L., Aguilar, G. R., Gray, A., ... & Group, A. R. L. (2022). *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis*. *The Lancet*, 399(10325), 629-655.
7. Cutting, E., Bordon, A., Khetan, S., Drouin, E., & Madan, S. (2024). *Forecasting the Fallout from AMR: Economic Impacts of Antimicrobial Resistance in Humans*. Center for Global Development.
8. B. Rittmann and M. P. Stevens, "Clinical Decision Support Systems and Their Role in Antibiotic Stewardship: a Systematic Review," *Current Infectious Disease Reports*, vol. 21, no. 8, p. 29, Jul. 2019, doi: 10.1007/S11908-019-0683-8.
9. *Ibid.*
10. K. Moulaei and K. Bahaadinbeig, "Diagnosing, Managing, and Controlling COVID-19 using Clinical Decision Support Systems: A Study to Introduce CDSS Applications," *Journal of biomedical physics & engineering*, vol. 12, no. 02, pp. 213-224, Apr. 2022, doi: 10.31661/jbpe.v0i0.2105-1336.
11. J. S. Albin et al., "Development and implementation of a clinical decision support system tool for the evaluation of suspected monkeypox infection," *Journal of the American Medical Informatics Association*, vol. 29, no. 12, pp. 2124-2127, Aug. 2022, doi: 10.1093/jamia/ocac151.
12. T.-H. Lin et al., "Artificial Intelligence-clinical Decision Support System for Enhanced Infectious Disease Management: Accelerating Ceftazidime-avibactam Resistance Detection in *Klebsiella pneumoniae*," *Journal of Infection and Public Health*, vol. 17, no. 10, p. 102541, Sep. 2024, doi: 10.1016/j.jiph.2024.102541.
13. S. M. Williamson and V. Prybutok, "Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare," *Appl. Sci.*, vol. 14, no. 2, p. 675, Jan. 2024, doi: 10.3390/app14020675.
14. A. Sakagianni et al., "Using Machine Learning to Predict Antimicrobial Resistance—A Literature Review," *Antibiotics*, vol. 12, no. 3, p. 452, Feb. 2023, doi: 10.3390/antibiotics12030452.
15. *Ibid.*
16. *Ibid.*
17. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in *KDD '16*. New York, NY, USA: Association for Computing Machinery, Aug. 2016, pp. 785-794. doi: 10.1145/2939672.2939785; R. Urena et al., "Proof of concept study on early forecasting of antimicrobial resistance in hospitalized patients using machine learning and simple bacterial ecology data," *Sci. Rep.*, vol. 14, no. 1, p. 22683, 2024.
18. R. Urena et al., "Proof of concept study on early forecasting of antimicrobial resistance in hospitalized patients using machine learning and simple bacterial ecology data," *Sci. Rep.*, vol. 14, no. 1, p. 22683, 2024;

T. K. Ho, "Random decision forests," in *Proceedings of 3rd international conference on document analysis and recognition, IEEE, 1995*, pp. 278–282. Accessed: Jul. 23, 2025. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/598994/>.

19. R. Urena et al., "Proof of concept study on early forecasting of antimicrobial resistance in hospitalized patients using machine learning and simple bacterial ecology data," *Sci. Rep.*, vol. 14, no. 1, p. 22683, 2024; "Keras: Deep Learning for humans." Accessed: Jul. 23, 2025. [Online]. Available: <https://keras.io/>; J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci.*, vol. 79, no. 8, pp. 2554–2558, Apr. 1982, doi: 10.1073/pnas.79.8.2554.

20. Williamson, S.M., & Prybutok, V. (2024). *Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare*. *Applied Sciences*, 14(2), 675.

21. *Ibid*; Elendu, C., et al. (2023). *Ethical Implications of AI and Robotics in Healthcare: A Review*. *Medicine*, 102(50), e36671.

22. Naik, N., et al. (2022). *Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?* *Frontiers in Surgery*, 9, 862322; Elendu, C., et al. (2023). *Ethical Implications of AI and Robotics in Healthcare: A Review*. *Medicine*, 102(50), e36671.

23. *Ibid*.



*Mr. Sophal Thear holds a master's degree in Artificial Intelligence for One Health (AI4OH) from Université Grenoble Alpes (UGA), France and a bachelor's degree in Computer Science from CADT. Currently, he is a Researcher and Lecturer in Data Science at Cambodia Academy of Digital Technology. He is also a member in a project: "Reducing Antimicrobial Resistance (AMR) via improving microbiology analysis capacity and optimizing the use of antibiotics through machine learning (ML) with AI.*

**Mr. Sophal Thear**



*Dr. Stanislas Rebaudet is an infectious diseases doctor at the European Hopital of Marseille and a public health researcher at SESSTIM (Aix-Marseille Univ, Inserm, IRD). He initiated this project to predict antibiotic resistance in hospitals, in order to help doctors to choose the right antibiotic therapy while waiting for the results of analyses carried out in the bacteriology laboratory.*

**Dr. Stanislas Rebaudet**



*Dr. Kennarey Seang holds a medical degree from the University of Health Sciences (UHS) Cambodia, MS in Nutrition from the National Institute of Public Health and a PhD in Epidemiology from the University of California Los Angeles (UCLA) in the USA. Since her PhD graduation in 2020, much of Dr Seang's work has focused on the epidemiology of HIV and social and behavioral aspects of people living with HIV. Dr Seang has led and co-led a number of projects, including a collaboration between the University of Health Sciences and the Kirby Institute at the University of New South Wales, looking at access to COVID-19 testing for PLWH in Cambodia, as well as community engagement research with the World Health Organization (Western Pacific Region) aiming to improve the access and retention rates in PLWH. Her current work involves using machine learning models with artificial intelligence to predict antimicrobial resistance and collaborations within the Exemplars in Global Health (EGH) program with the Hospital for Sick Children and the Gates Foundation on women's health and wellbeing (WHW) and maternal and infant nutrition and growth (MiNG).*

**Dr. Kennarey Seang**



*Dr. Sotharith Bory graduated from University of Health Sciences of Cambodia in 2007 for the General Medicine then continued for the specialist of Internal Medicine and graduated in 2010 at the same University. He continued his study in France where he completed his courses and received a diploma of DFMS (Specialize Medical Formation Diploma) in Internal Medicine in 2012 from University of Limoges, France Republic. In 2013, He was chosen to be an Infectious Diseases Physician at Calmette Hospital and since that time, he has been involved in many research and surveillance with other international researchers to conduct many research in the hospital and also the country. The main topic of his research career is Melioidosis and Antimicrobials Resistance. Beside working at the Hospital, Dr. Bory is also one of the lecturers for the Internal Medicine and the Infectious Diseases at the Faculty of Medicine, University of Health Sciences.*

**Dr. Sotharith Bory**

# HealthyRice: Enhancing Farmer-Expert Collaboration via Mobile Technology



Hongly Va and Soklong Him

## I. Introduction

In Cambodia, agriculture remains a key factor in the economy and society, supporting a significant part of the nation's workforce. According to the World Bank's collection of development indicators in 2022, 36.64% of Cambodia's employment was rooted in the agricultural sector.<sup>1</sup> The Cambodian government plays a crucial role in supporting the agricultural sector by encouraging investments and the diversification of products, along with improvements in irrigation and water management. As a result, agriculture continues to be a key component of Cambodia's economy, accounting for 22 percent of the nation's GDP in 2022 and providing jobs for 2.6 million individuals according to the Food and Agriculture Organization (FAO).<sup>2</sup>

For millions of farmers in Southeast Asia, rice serves as both a primary source of food and a crucial source of income. In 2024, Cambodia exported approximately 645,000 tons of processed rice.<sup>3</sup> The CRF announced that in the first four months of 2025, Cambodia exported 282,315 tons of milled rice, generating a total revenue of \$209 million.<sup>4</sup> Cambodia ranks as the 11<sup>th</sup> largest rice producer worldwide and is on track to reach a new high of 8.47 million metric tons (MT) (milled basis) for the 2024/25 marketing year.<sup>5</sup> As part of its strategic vision for 2025, CRF has set ambitious targets to increase its exports to one million tons.<sup>6</sup>

Cambodian rice has gained global recognition, having secured the esteemed World's Best Rice Award for the Phka Rumduol variety five times—in 2012, 2013, 2014, 2018, and 2022.<sup>7</sup> In 2023, it received the Golden Award for fragrant

rice from China.<sup>8</sup> The Cambodian government encourages sustainable farming practices and investments in rice processing technology to achieve the objective of enhancing productivity, improving quality standards, and ensuring long-term sustainability in the rice sector. This initiative is part of a broader effort to improve the competitiveness of Cambodian rice on the global market, focusing on enhancing quality, reducing costs, and ensuring environmental sustainability.

One of the primary challenges in rice cultivation in Cambodia is ineffective disease control and timely treatment. Among the most damaging issues are infestations by pests such as Brown planthopper (BPH) (*Nilaparvata lugens* (Stål)), Rice leaf folder (*Cnaphalocrociis medinalis*), Army worm (*Mythimna separata*), Thrips (*Stenchaetothrips biformis* (Bagnall)), Asian rice bug (*Leptocorispa* sp.), Leaf miner (*Hydrellia* sp.), Rice caseworm (*Paraponyx stagnalis*), Yellow stem borer (*Scirpophaga incertulas*), Striped stem borer (*Chilo suppressalis*) and Asian rice gall midge (*Orseolia oryzae*), which not only causes rice disease but also transmits destructive viruses.<sup>9</sup> Therefore, most farmers often rely heavily on pesticides to treat the rice disease. However, the massive use and misuse of pesticides can lead to environmental damage, pest resistance, and non-compliance with strict international export regulations. As a result, the Cambodian farmers struggle to meet the requirements of the international rice export market, especially in Europe, which applies strict rules on pesticides and requires labeling such as the eco label and environmental label.<sup>10</sup>

According to Ratha et al.,<sup>11</sup> the working age of the farmer falls between 31 and 64 years, representing a significant potential for adopting technologies and digital tools. However, this can lead to a major challenge due to the low level of digital literacy across the general population, including the agricultural sector.<sup>12</sup> Furthermore, the agricultural sector's technology that provides expertise and up-to-date information is still limited. Many farmers rely on traditional methods and community knowledge.<sup>13</sup>

Therefore, this chapter presents the mobile application HealthyRice, which has been developed to support rice disease monitoring and management through data collection, analysis, and information sharing among farmers, researchers, and agricultural experts. With the integration of AI technology, the application also explores the potential for automating disease recognition. This chapter focuses on Phase 1 of the project, which involves the design and development of the HealthyRice system.

## II. HealthyRice: An Overview of the Project

The HealthyRice project represents a major initiative and innovation designed to revolutionize Cambodia's rice cultivation practice by integrating digital technology with agroecological methods. The goal of the application is to enhance the rice plant's health, reduce pesticide usage, and improve the quality of rice for export. Consequently, the project development is divided into two stages.

The first stage is Mobile Development and Integration. In the project's first stage, we will focus on designing a user-friendly mobile application to assist rice farmers. The application will feature an intuitive user interface (UI) and user experience (UX) design that is approachable and easy to navigate for farmers with different levels of technological

skills. In addition, the application also comes with Khmer Automatic Speech Recognition (KhmerASR), allowing users to search for information using voice commands in Khmer, which makes it easier to access the resource without the need to type. Furthermore, Khmer Text-To-Speech (KhmerTTS) technology also integrates to read out the text contents to users, enabling users with limited literacy skills to engage and use the app without concerns. Last but not least, a comprehensive database of rice plant diseases and pests, as well as relevant information about the rice plant, was also constructed to store the information, which can be modified with up-to-date knowledge to maintain the rice crops.

The second stage is AI Model Development and Integration. In the project's second stage, we bring together Artificial Intelligence (AI) and Computer Vision to take the application to the next level. It initially starts with collecting a massive image dataset of rice disease and pests, since it requires a large amount of data to train, validate, and test the AI model that can recognize the disease accurately and consistently. In contrast, automated disease detection methods will simplify the process for the farmer to identify the disease manually, and early detection will help find a better treatment approach. Convolutional Neural Networks (CNNs) are primarily used for processing image data in various tasks of machine learning applications such as image recognition, image-based object detection, classification, etc. For this reason, we will utilize CNN and its well-known variants such as VGGNet, ResNet, DenseNet, MobileNet, and EfficientNet. The selected model will be trained end-to-end using labeled images of rice leaves affected by the specific disease occurring in Cambodia. To achieve high accuracy in the automated disease detection task, various experiments will be conducted to ensure the selection of a model that will be integrated into the applications.

### III. The HealthyRice Application

The main objective of HealthyRice is to provide farmers with a simple and powerful cross-platform mobile application to learn and rely on as a virtual technical expert for identifying the disease in the rice crop. In this section, we will present the key innovative and involved technology used in the proposed application.

#### 1. HealthyRice: A Mobile Application

The mobile application was developed using Flutter since it is an open-source software development toolkit supported by Google. The application allows users to view treatments based on the specific disease that they are looking for and the pathogens damaging their rice, report new disease cases, and contribute to a broader agricultural community. Importantly, our proposed application supports both English and Khmer languages to enhance accessibility for its main users. Figure 1 shows a use-case diagram of how the farmer can interact with the proposed application.

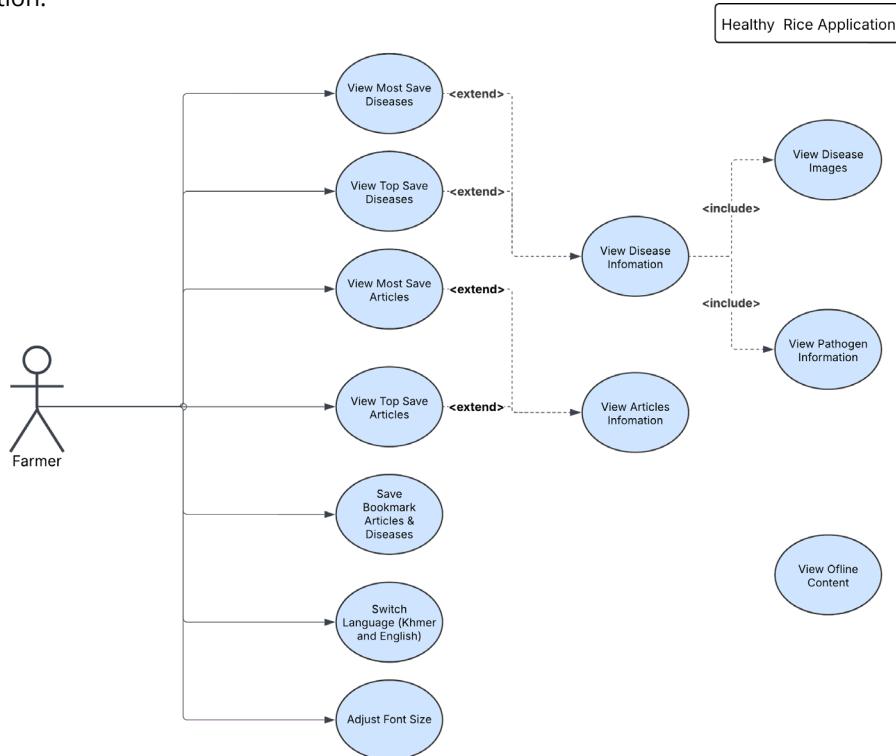


Figure 1: Use Case Diagram

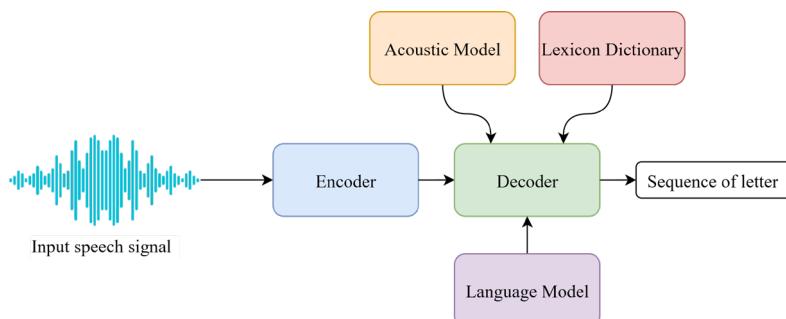
To make the application easy to use and enhance the user experience, our application does not require user authentication to allow instant access. Similarly, due to the insufficient internet infrastructure across Cambodia, our project includes offline capabilities in order to support all of the farming areas where internet penetration is notably lower. Table 1 describes the details of functional requirements, which are the core functionality needed to ensure usability and accessibility.

Table 1: Functional Requirements Table

Feature	Description
Details of Rice Disease	View detailed information on rice diseases, including the main causes, pests, symptoms, associated pathogens, and recommended treatments.
Most Common Diseases (Monthly Trends)	Display a list of the most frequently viewed diseases each month, based on user interactions (the clicks on disease details of other users).
Top Saved Diseases	Shows a list of diseases that users have saved for quick reference.
Most Viewed Article	Highlights the most frequently viewed articles, which are provided by administrators and cover topics such as disease prevention and treatment.
Top Saved Article	Save articles for later reference, and this feature displays the most saved articles.
ASR for Search	Perform voice-based searches by converting spoken words to text. This is particularly useful for searching disease information using local or technical terms.
TTS Functionality	Reads content aloud, such as rice disease details and treatment information, in the Khmer language.
Reporting	Report a new disease that is not found in the existing category.

Additionally, Khmer-ASR technology for voice-based disease searches features allow users to engage with the content and articles provided by administrators. Figure 2 shows the architecture of Khmer-ASR integrated in the application. The ASR system processes spoken audio by first extracting acoustic features from the input signal. These features are then interpreted by the acoustic, language, and pronunciation models within the decoder to generate the most probable sequence of words corresponding to the spoken input.<sup>14</sup>

- An acoustic model that is responsible for recognizing the acoustic features and phonemes (the smallest unit of sound in language).
- Language model response for computing the probability of a word sequence in a given language.
- Lexicon dictionary response for mapping words to phonemes

Figure 2:  
The Khmer-ASR Architecture

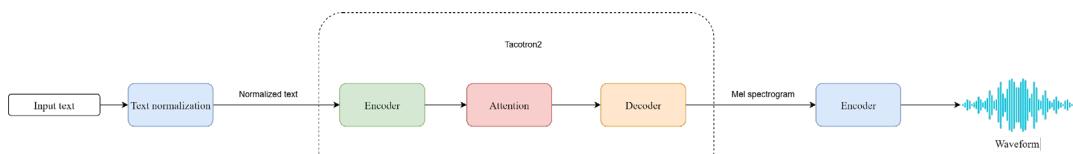


Figure 3: The Khmer-TTS Architecture

In addition, we also integrate the Khmer TTS to provide accessibility for users with limited literacy skills. The Khmer TTS system first takes the text as input and converts the written text into spoken text or a list of symbols using text normalization. The encoder encodes the text and generates the spectrogram.<sup>15</sup> After that vocoder takes the spectrogram and generates the waveform. Figure 3 shows the architecture of Khmer TTS.

## 2. HealthyRice: An Administration Portal (Web-Based)

Besides the mobile application, we have also developed a comprehensive, web-based management portal to centralize data and strengthen the collaboration between farmers and agricultural experts. The web-based portal is primarily designed for use by researchers, agronomical experts, and administrators, who are in charge of driving agricultural development and decision-making in Cambodia. The primary functional features of the portal are:

- *Application Programming Interface (API) Management:* The administrator manages the core data content to display in the mobile application. The main categories are: the category of the disease and its detailed information, its primary causes (e.g., environmental, nutritional), associated pests, observable symptoms, specific pathogens, and expert-recommended treatments.
- *Article Management:* The administrator can write and publish the articles to the mobile application's users. The article can provide up-to-date information on

best practices, preventative measures, new research findings related to rice cultivation, and seasonal agricultural advice. To align with the experts' goal, the comprehensive resources on Integrated Pest Management (IPM) strategies for rice cultivation can also be prepared and published for the users.

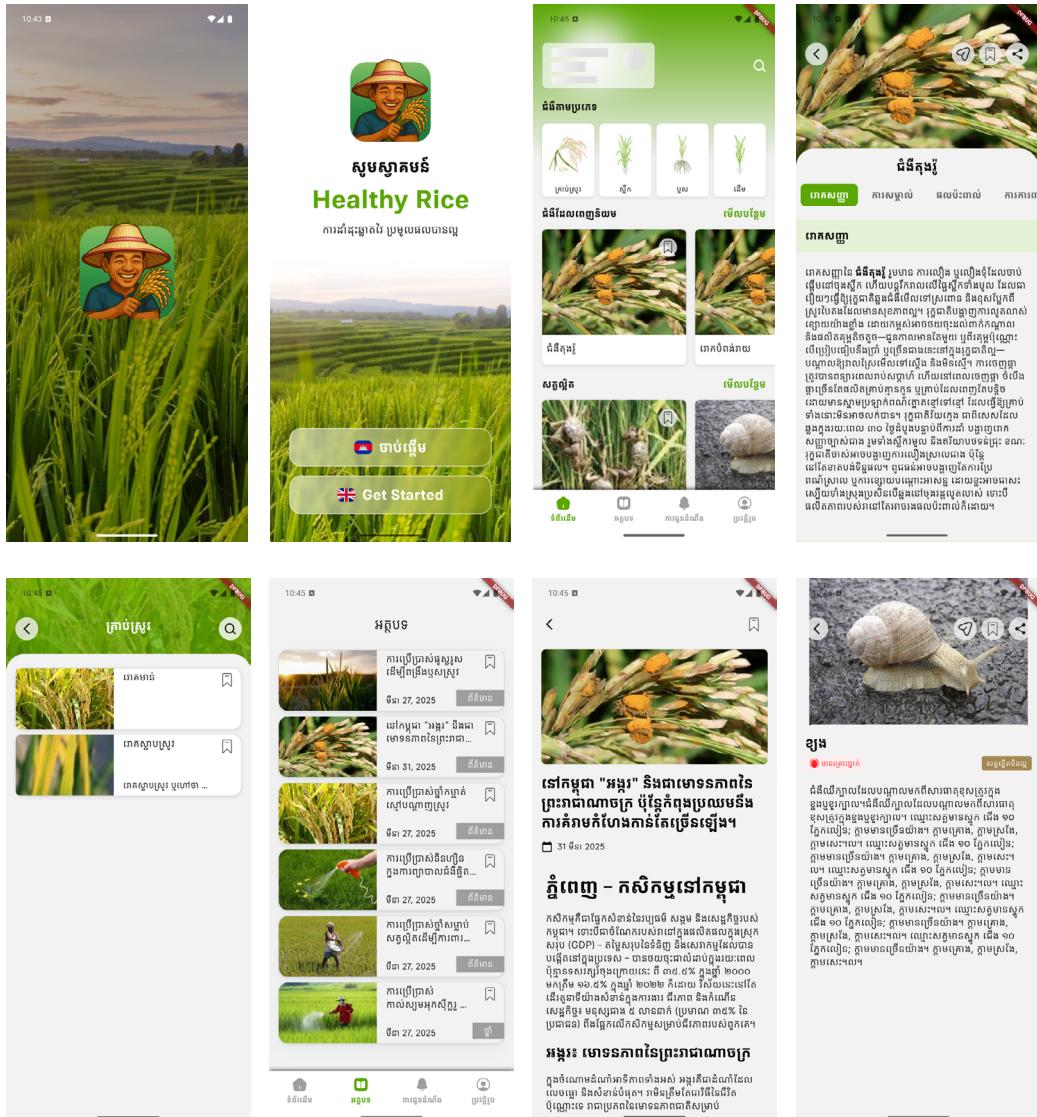
- *Notification Management:* The administrators can send a notification to a target group of users with the latest information or important information directly to the HealthyRice mobile application. These features are used for several useful purposes such as notifying the users about potential disease outbreaks in their area, offering timely agricultural advice, sharing new content which is available on the app, and delivering urgent public health messages related to farming.
- *Report Review and Analytics:* The administrator can review and validate the reported disease submitted by the user through the mobile application. This feature allows the expert to identify the trend of such an outbreak from a specific area or location of the reported case.

## IV. Results and Discussion

The HealthyRice application comes with a mobile application for the farmers and a web-based administration portal for experts in rice that includes researchers and agronomists.

## 1. Mobile Applications

The completed elements of the mobile application consist of several important screens: the Splash Screen, which provides the first visual introduction to the app; the Main Screen, allowing users to access different features; the Article Screen, where users can view the content; and the Profile Screen, where users can adjust their personal information and settings.



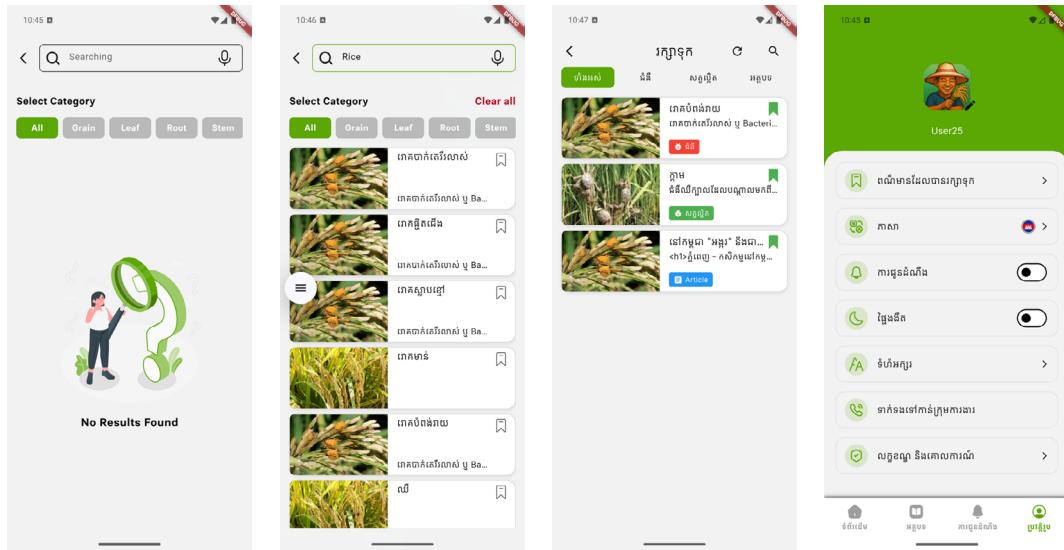


Figure 4: Key User Interface Snapshots of the HealthyRice Mobile Application

## 2. Web Portal

The web portal for administrators has been developed and interacts with the Mobile Application seamlessly through the design API. Unlike users on the mobile application, the admin side requires user authentication to effectively secure content and information. A user-friendly dashboard will be displayed on the homepage of the portal. Therefore, the experts and researchers as admin users can manage content, publish articles on best practices or IPM, and send timely notifications to farmers through this platform.

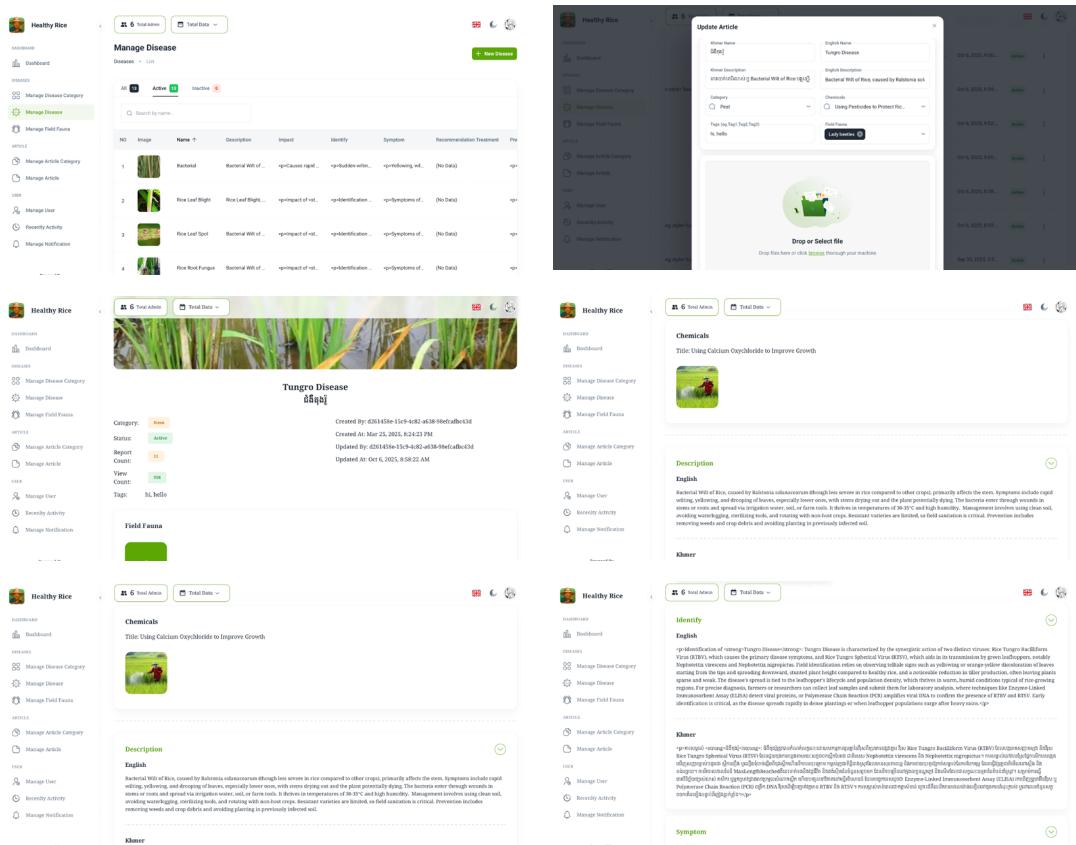


Figure 5: Key User Interface Snapshots of the HealthyRice Web Portal for Administrator

### 3. Challenges and Discussions

The proposed application aims to support the Cambodian farmers in enhancing and improving their rice cultivation knowledge through both a mobile application and a web portal. However, the developed application comes with several challenges that need to be addressed. First, it is important that after the development phase, the design pilot testing on a small group of people should be done in order to receive feedback and improve the new design and feature. Similarly, the issue of digital literacy can also impact the usage of our application; therefore, an intensive training session is conducted regularly to make the farmers familiar with our system. Moreover, the administrator needs ongoing content updates to ensure that the information provided is relevant and reflects the latest research and agricultural practices. The new collaborator, such as agricultural experts from

government agencies and local universities specializing in agriculture, can significantly curate accurate and timely content.

## V. Conclusion

The HealthyRice project presents a practical solution to connect Cambodian rice farmers with agricultural experts through a proposed digital platform that includes a mobile application and a web-based portal. To speed up the development process, we have integrated Khmer ASR and KhmerTTS, which were developed in-house by our Institution (CADT), to ensure accessibility for farmers with different literacy levels. Additionally, offline functionality addresses the challenges of internet connectivity in rural areas. Although the project's first phase has been completed, pilot testing is necessary to gather user feedback on the proposed application. The

pilot testing should involve a small group of farmers, relevant government staff, and an agricultural expert in order to improve the application before the official release.

In future work, the project's second phase can be initiated by analyzing the most common rice diseases and pests. Consequently, a data collection strategy should be implemented to obtain high-quality image data to develop a robust and accurate AI-based disease detection model. Both primary and secondary data sources should be obtained to increase the variability of the disease images.

## **VI. Acknowledgement**

I would like to express my sincere gratitude to Dr. Stéphane BELLAIORE and Dr. Rottana Ly for providing the resources and support necessary to complete this chapter. Special thanks to CADT for their generous support, which made this work possible.

## References

1. "Cambodia Employment in Agriculture - Data, Chart." n.d. *TheGlobalEconomy.com*. [https://www.theglobaleconomy.com/Cambodia/Employment\\_in\\_agriculture/](https://www.theglobaleconomy.com/Cambodia/Employment_in_agriculture/)
2. "Cambodia Investment Proposal." 2024. *HandInHand*. 2024. <https://www.fao.org/hand-in-hand/previous-editions/hih-IF-2024/cambodia/en>.
3. "Cambodia Rice Federation." 2023. *Crif.org.kh*. 2023. <http://www.crif.org.kh/>.
4. "Cambodia Earns 209 Mln USD from Milled Rice Export in First 4 Months." 2025. *News.cn*. 2025. <https://english.news.cn/20250505/0ce3becb2b6c46ef8a56c66>
5. "Foreign Agricultural Service Global Market Analysis International Production Assessment Division." 2025.
6. "Cambodia Rice Federation." 2023. *Crif.org.kh*. 2023. <http://www.crif.org.kh/>.
7. "The 7th Cambodia Rice Forum 2024." 2023. *Crif.org.kh*. 2023. [http://crif.org.kh/?page=api\\_location\\_detail&id=1572&lg=en](http://crif.org.kh/?page=api_location_detail&id=1572&lg=en).
8. afid. n.d. "Cambodia's Fragrant Rice Won the Gold Award in the 1st China-ASEAN Taste Quality Appraisal Activities of High-Quality Indica Rice Variety in China." AFID. <https://afid.cdc.gov.kh/news/fragrant-rice/>.
9. Dunn, Lucinda, Tanya Latty, van Ogtrop, and Daniel. 2023. "Cambodian Rice Farmers' Knowledge, Attitudes, and Practices (KAPs) Regarding Insect Pest Management and Pesticide Use." *International Journal of Agricultural Sustainability* 21 (1). <https://doi.org/10.1080/14735903.2023.2178804>; Sievers, Benjamin L., Sudipta Hyder, Filip Claes, and Erik A. Karlsson. 2024. "Ingrained: Rice Farming and the Risk of Zoonotic Spillover, Examples from Cambodia." *One Health* 18 (June): 100696. <https://doi.org/10.1016/j.onehlt.2024.100696>; Kakada Oeum, Malyna Suong, Kimsrong Uon, Léa Jobert, Stéphane Bellafiore, Aurore Comte, Emilie Thomas, Fidoro Kuok, and Lionel Moulin. 2024. "Comparison of Plant Microbiota in Diseased and Healthy Rice Reveals Methylobacteria as Health Signatures with Biocontrol Capabilities." *Frontiers in Plant Science* 15 (October). <https://doi.org/10.3389/fpls.2024.1468192>.
10. Horn Sarun et al., "Challenges and Opportunities in Rice Cultivation and Marketing of Cambodia: A Review Article," *International Journal of Innovative Science and Research Technology* 10, no. 1 (January 2025): 1, <https://doi.org/10.5281/zenodo.14862890>.
11. Hem, Ratha, Makararavy Ty, Asikin Yoeu, Rylida Vong, Luis Molina Carpio, and Siphann Hoem. 2024. "UNLOCKING the DIGITAL POTENTIAL: EVALUATING SMARTPHONE-BASED DIGITAL COMPETENCE among CAMBODIAN FARMERS." *International Journal of Applied Science and Research* 07 (04): 28–40. <https://doi.org/10.56293/jasr.2024.6004>.
12. Natasha Beschorner et al., *Benefiting from the Digital Economy: Cambodia Policy Note* (Washington, D.C.: World Bank Group, 2018), <http://documents.worldbank.org/>

curated/en/100841543598854492.

13. "Traditional Rice Farming." n.d. [www.gvi.co.uk](http://www.gvi.co.uk). <https://www.gvi.co.uk/blog/traditional-rice-farming/>.
14. Srun, Nalin, Sotheara Leang, Ye Kyaw Thu, and Sethserey Sam. "Convolutional time delay neural network for Khmer automatic speech recognition." In 2022 17th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), 1-6. IEEE, 2022. <https://doi.org/10.1109/iSAI-NLP56921.2022.9960286>.
15. "Text-To-Speech with Tacotron2 — Torchaudio 2.7.0 Documentation." 2025. Pytorch.org. 2025. [https://docs.pytorch.org/audio/stable/tutorials/tacotron2\\_pipeline\\_tutorial.html](https://docs.pytorch.org/audio/stable/tutorials/tacotron2_pipeline_tutorial.html).

**Mr. Hongly Va**

*Hongly Va received a B.E. degree in Information Technology Engineering from the Royal University of Phnom Penh in 2019. He received his Ph.D. in Software Convergence from Soonchunhyang University, Asan, Republic of Korea, in 2024. Currently, He is a senior lecturer and researcher of CADT. His research interests include Computer Graphics, VR, Parallel Computing, Physically-based Modeling and Simulation, and Computer Vision.*

**Mr. Soklong Him**

*Soklong Him received a B.S. degree in Computer Science from CADT, Phnom Penh, Cambodia, in 2021. He received his M.S. degree in Natural Language Processing (NLP) from the University of Lorraine, Nancy, France, in 2023. Currently, he is a lecturer and researcher of computer science at CADT. His research interests include Khmer NLP, Automatic Speech Recognition(ASR), Text-To-Speech (TTS), speech anonymization.*

# Policy-Driven Approaches in AI-Powered Verification for Combating Mis- and disinformation: A Cambodia-Focused Study



*Prohim Tam, and Riccardo Corrado*

## I. Introduction

Cambodia's digital ecosystem can be indicated by a notable paradox: high digital connectivity paired with low digital literacy. With fast-growing mobile-broadband subscriptions reaching 101.97% in 2022 (as shown in a study of the World Bank "Digital Progress and Trends Report 2024" for the East Asia and Pacific region),<sup>1</sup> of the overall population of 17,201,724 people in 2022,<sup>2</sup> and a remarkably high percentage of citizens on social media platforms, the country has a strong foundation for digital engagement. However, a significant gap remains in the population's ability to effectively leverage these technologies to maximize the benefits and not expose them to digital vulnerabilities. Studies from 2020 indicate that only 32% of Cambodians possess basic digital access, with just 5% having intermediate skills and only 1% demonstrating advanced proficiency.<sup>3</sup> This imbalance highlights a critical challenge in ethical digital adoption and showcases difficulties in realizing the full potential of the country's digital transformation initiatives.

In light of these challenges, Cambodia initiated the development of an extensive policy framework in 2020 aimed at steering its digital, economic, and social advancements. A significant aspect of this initiative is the enhancement of Digital, Media, and Information Literacy, abbreviated as DMIL.<sup>4</sup> The primary goal of this competency framework is to move beyond mere digital access and cultivate digital citizens who possess the necessary

skills to navigate the digital landscape securely and efficiently.

Digital literacy is the ability to understand and use digital platforms in a safe, ethical, responsible, and productive manner for accessing information, learning, comprehending, evaluating, and analyzing, synthesizing information in the online sphere, including how AI algorithms generate content (video, audio, text, and images). Such understanding is necessary for protecting the inclusivity of growth, and also to limit dangers from biased and fake information produced by AI. In 2022, smartphone connectivity reached 82.6%,<sup>5</sup> offering an increasing number of users the opportunity to be connected and informed, but also, on the other hand, being possible victims of AI-generated content, produced with the scope to mislead or manipulate. Figure 1, from the World Bank digital report released in 2024, shows where we are in the East Asia and Pacific region when it comes to types of phone connections. Comprehensive approaches, in the form of formal education programs, training workshops, and seminars, should be implemented across different sectors to address these issues effectively.

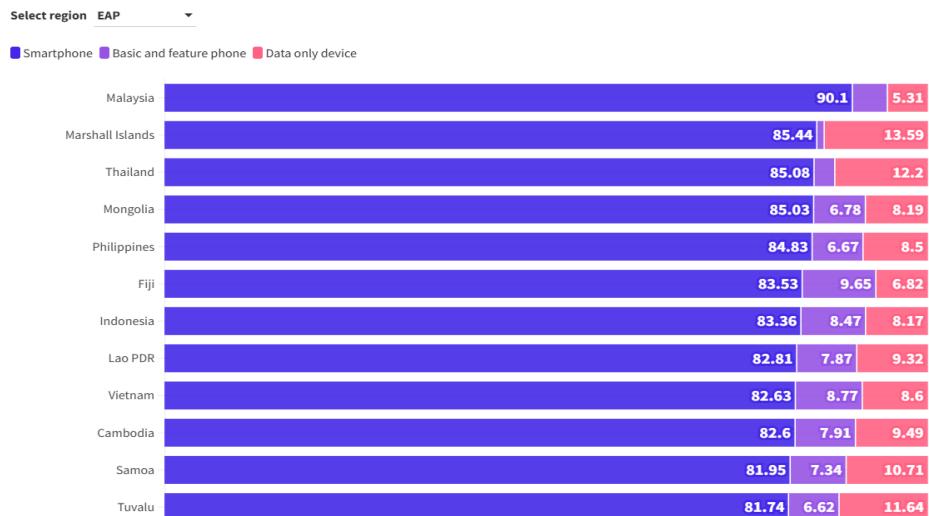


Figure 2 - Mobile phone connections by type 2022

Cambodia is facing a low digital literacy problem, as highlighted also by Pa Chanroeun, the president of the Cambodia Institute for Democracy, who at the 2024 Phnom Penh Internet Forum noted how many Cambodians still lack essential digital literacy skills, making them vulnerable to online threats, misinformation, and disinformation.<sup>6</sup> From the Digital Literacy for Employability report by the United Nations Development Programme (UNDP), it was highlighted that at the current level of literacy, which has improved over the years across Cambodia, youth can be expected to be capable of searching, identifying and retrieving information; however the ability to filter and evaluate information is still restricted. The low digital literacy problems,<sup>7</sup> expose the Cambodian population to several significant risks in today's digital age, including:

- The massive spread of mis- and disinformation: Disinformation actions can be highly destructive to public trust and threaten social stability. Between 2017 and 2022, Cambodia's Ministry of Information reported approximately 4,000 cases of disinformation.<sup>8</sup> Recently, the emergence of AI-generated images depicting individuals with disabilities soliciting donations has begun to create significant problems within Cambodian society.<sup>9</sup> Unfortunately, many individuals continue to be deceived by these scams.
- Risks from emerging technologies: The exponential growth of technology adoption, such as AI, poses new challenges, especially to the rise of content creators. Technology often contains embedded biases and addictive algorithms that can lead to excessive use, misinformation, and privacy violations; furthermore, AI-generated content itself requires the content creators to deeply understand intellectual property rights, acknowledging the importance of trademark and avoiding the creation and dissemination of harmful or misleading content for their own clout.
- Digital security: Concerns about issues of identity theft, privacy breaches, and digital security are widespread, showcasing the need for citizens to acquire the knowledge necessary to protect themselves in cyberspace.

With these abovementioned social challenges, the Information Ministry's 2025 campaign titled "Say No to Fake News" has been launched. From 2023 to 2024, the number of recorded fake news cases increased by 14%, increasing from 3,208 to 3,651.<sup>10</sup> Most of the targeted incidents

were pointed to national security, public order, politics, diplomacy, military affairs, and government leadership.<sup>11</sup> A press conference declared AI-driven misinformation a national security concern, while promoting the newly issued Charter for Professional Journalism (2024) and commitments to ethical reporting.<sup>12</sup>

MPTC has been actively working to enhance capacity building on digital skills, a priority in the draft National AI Strategy for 2025-2026. The strategy outlines 10 principles and 4 actions, which demonstrate a serious commitment to improve human resources, including management level, experts, officials, teachers, and citizens, as well as combat AI-driven mis- and disinformation.<sup>13</sup>

In the light of the discussion above, this chapter aims to offer:

- A systematic review of state-of-the-art AI-powered applications for mis- and disinformation detection in low-resource languages.
- A proposed framework prototype combining FastText and TF-IDF to detect disinformation in Khmer texts.
- A policy roadmap and stakeholder engagement aligned with the Cambodia Competency Framework, which emphasizes transparency and rights, to make the development of AI-powered detection possible for combating against mis- and disinformation.

The paper structure is organized as follows. Section II presents the literature review of AI detection methods. Section III presents the design of the hybrid verification framework, which combines FastText and TF-IDF. Section IV outlines the policy recommendations for relevant stakeholders. And finally, section V concludes the paper.

## II. Literature Review

AI-powered applications for mis- and disinformation detection can be categorized based on the integrated machine/deep learning algorithms that analyze different features of the information, including content, social context, and propagation patterns. In our literature review, we query the researcher hubs such as Google Scholar and Semantic Scholar to find relevant studies for low-resource languages, and we categorize the algorithm choices in fake news detection as follows:

### 1. Term Frequency-Inverse Document Frequency (TF-IDF) with Traditional ML

Both Bag-of-Word (BoW) and TF-IDF methods were used to transform text into feature vectors.<sup>14</sup> Additionally, character and word n-grams (up to 10-grams) were explored. Multiple classifiers were trained with optimized hyperparameters. Deep learning models (MLP and LSTM) were also implemented. Specifically, the author manually compiled a dataset of 900 Urdu news articles, balanced between 500 real and 400 fake articles, spanning five domains: business, health, entertainment, sports, and technology. The dataset was divided into 75% training and 25% testing sets. Finally, the performance was assessed using accuracy, precision, recall, F1 score, and confusion matrix components. Cross-validation was performed to ensure robustness.

Authors introduced a hybrid feature extraction method called Hybrid TF-IDF, which combines word embedding (Word2Vec) to capture semantic and syntactic features.<sup>15</sup> Ensemble learning, specifically the AdaBoost classifier, is employed to improve classification performance by combining multiple weak classifiers such as ID3, Random Forest, and Naive Bayes.

Authors utilized the WELFake dataset comprising 72,134 news articles (35,028

real and 37,106 fake), sourced from multiple datasets (e.g., McIntyre, Kaggle, Buzzfeed Political, Reuters). For feature extraction, the system employed a hybrid TF-IDF (combining Word2Vec with TF-IDF) to generate weighted vectors representing contextual semantics. The study implemented an AdaBoost ensemble classifier that combines ID3, RF, and Naive Bayes classifiers using a voting mechanism.

## 2. FastText Embeddings for Subword Morphology and Cross-Lingual Semantics

TF-IDF highlights discriminative words based on frequency and rarity. However, it requires a deep understanding of the cultural context in each aspect of the language. FastText extends Word2Vec by considering subword information, which improves handling of rare words and Convolutional Neural Networks (CNNs) designed to automatically learn hierarchical features from text data.<sup>16</sup> For data collection, the study utilized the TruthSeeker dataset, which comprises over 180,000 labeled tweets and social media posts from 2009 to 2022, with multi-layered validation including expert review and crowdsourcing.<sup>17</sup> For pre-processing, the procedures included text cleaning (removing URLs, special characters, and emojis), tokenization, stop word removal, and splitting data into training (70%) and testing (30%) sets. In embedding techniques, that applied TF-IDF was deployed with trained Word2Vec and FastText models on the training data, and then the training and testing datasets were vectorized accordingly. Three CNN architectures with variations in convolutional layers, pooling strategies, and dropout mechanisms are used.

## 3. Cross-Lingual Deep-Learning Model on Pre-Trained BERT

A system trained in a high-resource language (e.g., English) can be generalized for fake news detection to a low-resource language (e.g., Khmer) using bilingual alignment and joint encoding. The process of identifying fake news

across multiple languages by transferring knowledge from a resource-rich language to a low-resource language can be achieved by utilizing pre-trained language models such as BERT and multilingual-BERT to encode news content.<sup>18</sup> Data annotation was conducted by manual collection and labeling of Chinese COVID-19 news based on existing English fact-checking datasets. For long texts, the procedure of breaking lengthy news articles into sub-texts is mandatory to accommodate model input limitations and retain content integrity.

For data collection methods, authors curated a Chinese dataset with 86 fake and 114 real news articles, based on fact-checking sources and manual annotation. The process included translating Chinese news into English for model input, which ensures alignment with English training data. The analysis techniques employed a deep learning framework named CrossFake, built on pre-trained BERT models. Furthermore, encoded sub-texts individually, were aggregated embeddings via mean pooling, and classified credibility. The evaluation compared baseline models of monolinguals.

## 4. Hybrid Vectorization Techniques

The critical issue of fake news detection in the context of the expanding influence of social media and online content dissemination is severe. Hybrid techniques tend to improve detection accuracy, efficiency, and adaptability, which contribute to more reliable tools against misinformation with societal significance. Multiple datasets, including FA-KES (conflict-related news), WELFake, Kaggle datasets, and a merged dataset with diverse news sources, are used.<sup>19</sup> For feature extraction, the following approaches are deployed: 1) TF-IDF for statistical word importance, 2) Empath for semantic and psychological features, and 3) A Hybrid approach combining both for richer feature sets.

As a result, the hybrid feature extraction approach outperformed individual methods, achieving up to 99% accuracy with gradient boosting on large datasets. Logistic regression proved most reliable in overall prediction accuracy, while Naive Bayes models (MNB and BNB) offered the fastest runtimes (approx. 11 seconds) but varied in accuracy depending on dataset size. Transfer learning improved computational efficiency but showed mixed results in accuracy depending on dataset alignment. Smaller datasets such as FA-KES yielded lower accuracy (approx. 45–57%), which indicates challenges with limited data. As an overall recommendation, combining linguistic, psychological, and morphological features with statistical methods enhances detection robustness, as well as improves nuance in feature space and reduces false positives in emotionally charged misinformation cases.

### III. Proposed Framework

Following the literature review, we identified major challenges for a Cambodia-focused study, including data scarcity, lack of context, processing long structured text, generalizability, and balancing performance with efficiency. In our proposed framework, we provide an overview of recommended settings and tools to tackle the above-mentioned challenges, following our review prototype for Khmer language detection.

#### 1. Model Selection

The employed feature extraction methods in this chapter are TF-IDF and FastText, due to their computational efficiency and, mostly, their ability to work well also with a limited dataset, a condition fitting the Cambodian paradigm. Both methods offer a unique advantage, especially for low-resource and morphologically rich languages such as Khmer.

FastText creates “word vectors” that are high-dimensional numerical representations of words that reflect their meaning. It is an

extension and the latest version of traditional CBOW with the Skip-gram model that uses subword information. It breaks words into bags of character n-grams, which allows the model to consider both morphemes and subword units. This feature specifically allows FastText to be powerful with languages that are morphologically complex and to deal properly with out-of-vocabulary words at training time. FastText improves the Word2Vec approach by including subword information. The process starts with data preprocessing, which involves applying multiple operations to get the dataset in a proper, thoroughly clean format so that it is used for training. It considers Khmer punctuations, special characters, spaces, numbers, and also English words, stop word removal, and word tokenization. In Figure 2, we represent the data pre-processing.

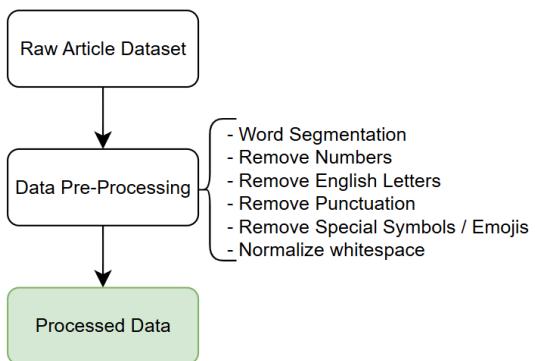
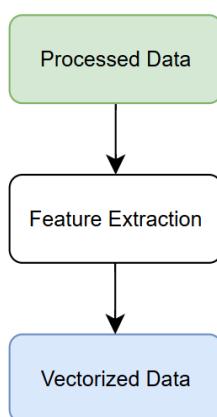


Figure 2 – Data Pre-Processing for Khmer Text

Next, the processed data are put through each of the feature extractions TF-IDF and FastText as illustrated in Figure 3 is to capture relevant features and convert those features into meaningful representations called vectorized representation of data.

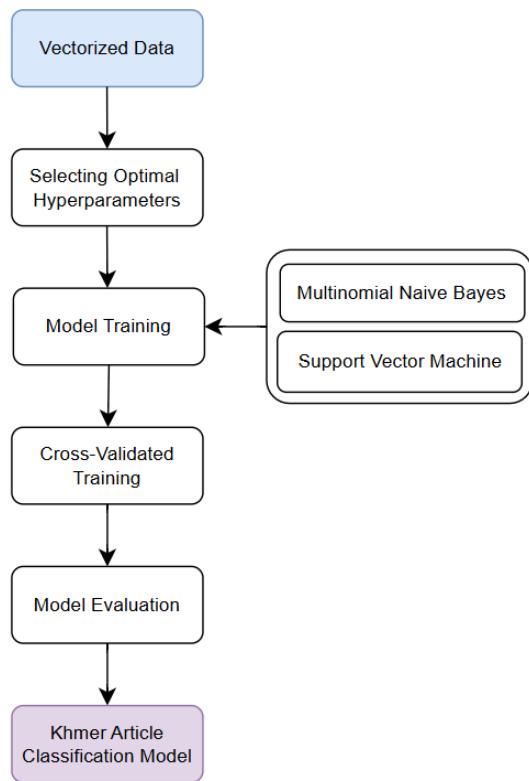


*Figure 3 - Text Vectorization and Feature Extraction Using TF-IDF and FastText*

To ensure the validity of model training, optimal hyperparameters are set using the fixed parameter, which provides the same threshold to do comparative performance analysis on the feature extraction TF-IDF and FastText. The following are the optimal parameters used for model training:

- MNB with its default parameters
- SVM with  $C=0.1$ ; `decision_function_shape: "ovr"`; `gamma: scale`; and `kernel= "rbf"`

After getting the optimal hyperparameters, the model will then undergo k-fold validation ( $k=10$ ), where  $k$  denotes the number of groups or folds in a given data sample. The cross-validation will help verify the classification result, ensuring an accurate numerical output as illustrated in Figure 4.



*Figure 4 - Selected Machine Learning Model in the Proposed Framework*

## 2. Model Evaluation

In regard to the model evaluation, the evaluation metrics used include accuracy, precision, recall, and F1-score. For the statistical analysis, results from the evaluation metrics, including accuracy, precision, recall, and F1-score for each classifier model training, will be used. Descriptive statistics will be employed to find the mean, standard deviation, and variance in order to assess the impact of feature extraction on the performance of both classifier models. Furthermore, to determine the best model performance between TF-IDF and FastText with the combination of both classifier models, a paired t-test will be performed to identify which feature extraction offers a better multi-class classification performance towards fake news detection (compared to the real news article).

### 3. Application Development

For front-end development, the application can be implemented using the Streamlit framework, a Python-based tool that enables the rapid development of user-friendly interfaces for machine learning applications. The back-end integrates a classification model that processes user input and returns a category prediction in real time. The entire application, including both the front-end and back-end components, can be hosted on Streamlit, ensuring accessibility and ease of deployment. Table 1 shows the parameters and specifications.

The functional workflow of the prototype is structured as follows:

- User Input: The user uploads a Khmer news article file through the web interface.
- Request Processing: The uploaded file is sent to the server for analysis.
- Model Classification: The server processes the request, analyzes the article's content, and classifies it into the most relevant category; then labels the real or fake content.
- Response Display: The categorized result, along with the document title, is displayed on the application's interface for the user.

*Table 1 - Specification for Application Development*

Specification	Description
Deployment	Streamlit Cloud's free tier (1GB RAM, 1 CPU core).
Language	Python 3.8+
Streamlit	≥ 1.28
Scikit-learn	≥ 1.2
FastText	≥ 0.9.4
Khmer-nltk	≥ 1.0

## IV. Policy Recommendation

This chapter aimed to partially address questions such as *how Cambodia can begin pushing for AI-verification and detection tools to combat mis- and disinformation, and which are key policies that should be implemented or introduced to speed this up?* Our analysis provides a possible coordinated policy framework suitable for Cambodia's ecosystem. The adoption of the proposed framework should interest five key stakeholder groups:

### Government Sector

- Ministries should establish and publicly maintain a verified register of official social media profiles and websites (registered verifiable news media) so that citizens can easily distinguish authentic sources from unverified or misleading ones.

- Open public information should be promoted by requiring targeted publicly funded institutions to publish sanitized, machine-readable datasets as part of their regulatory work and by adopting open-data norms that strengthen AI development through higher-quality, more robust datasets.

### Local News Media

Cambodian media must evolve from passive publishers to active verification stakeholders:

- *Charter-Compliant Verification Desks:* Establish a verification bureau within leading Khmer-speaking publications trained to operate AI flagging and human moderation hybrid models.
- *Trusted Data Sharing:* Facilitate the release of authoritative journalism archives, timestamps and metadata to academic and NGO partners for model training in a manner that safeguards user privacy and national security standards through data-use agreements.

### Academic Institutions

Universities and research centers serve as the long-term backbone for technical and linguistic capacity building:

- *Interdisciplinary Collaboration:* Propose hackathons or workshops, in which students from different majors (e.g., IT, policy studies) can come together to iterate on prototypes.

### Technology Developers

Private-sector and independent developers must adopt high standards to guarantee transparency, trust, and contextual relevance:

- *Adopt Ethical AI Principles:* Publish datasheets for Khmer models, in alignment with global norms (UNESCO values of fairness, non-discrimination, privacy, explainability, etc.)
- *Continuous Performance Reporting:* Commit to biannual public release of performance audits across metrics (precision, recall, false-positive rates) and updating fallback protocols for mis-flagging errors.

## V. Conclusion and Future Work

The hybrid FastText and TF-IDF architecture we proposed offers possibilities for local developers with an approach that is both technically robust and lightweight enough to prototype and deploy using low-cost web tools, drastically reducing development overhead and enabling rapid iteration even in resource-constrained settings. Critically, by pairing this model with open-data policies from government agencies, mandating the release of verifiable, machine-readable sources such as official press statements, speeches, and local media archives, developers can gain access to high-quality, locale-specific training data capable of grounding their models in the real Cambodian context.

Next steps may include implementing the framework in a complete detection and information sharing network to establish the viability of gleaning viral mis- and disinformation based on graph-based analysis, as well as conducting user behavior studies and measuring resilience to be able to test educational interventions at an individual level in order to explore how literacy efforts may affect reader behaviors and attitudes toward false content.

## References

1. World Bank Group, "Digital Progress and Trends Report: Interactive Charts," World Bank Group, March 4, 2024, accessed August 1, 2025, <https://www.worldbank.org/en/data/interactive/2024/03/04/digital-progress-and-trends-report-interactive-charts>.
2. World Bank Group, "Population, Total – Cambodia," World Bank Data, accessed August 1, 2025, <https://data.worldbank.org/indicator/SP.POP.TOTL?locations=KH>.
3. UNESCO, "Cambodia Launches Its First Competency Framework on Digital, Media and Information Literacy to Empower Citizens in Today's Digital Society," UNESCO, July 2, 2024 (last updated July 4, 2024), accessed August 1, 2025, <https://www.unesco.org/en/articles/cambodia-launches-its-first-competency-framework-digital-media-and-information-literacy-empower>.
4. Ministry of Post and Telecommunications, (Cambodia), "Digital, Media and Information Literacy (DMIL) Competency Framework (English version, PDF)", June 2024, accessed August 2, 2025, <https://www.khmersme.gov.kh/wp-content/uploads/2024/06/dmil-english.pdf>.
5. World Bank Group, "Digital Progress and Trends Report: Interactive Charts," World Bank Group, March 4, 2024, accessed August 1, 2025, <https://www.worldbank.org/en/data/interactive/2024/03/04/digital-progress-and-trends-report-interactive-charts>.
6. Seoung Nimol and Tang Porgech, "Phnom Penh Internet Forum Calls Attention to Digital Illiteracy and the Fight for Internet Freedoms," CamboJA News, October 18, 2024, accessed August 2, 2025, <https://cambojanews.com/phnom-penh-internet-forum-calls-attention-to-digital-illiteracy-and-the-fight-for-internet-freedoms/>.
7. United Nations Development Programme (UNDP), "Digital Literacy for Employability and Entrepreneurship Among Cambodian Youth: Assessment Report," by Mr. Nimpl Vamoeurn, Phnom Penh, September 2020, accessed August 2, 2025, <https://online.anyflip.com/qcffd/bhfx/mobile/>.
8. Ministry of Information (Cambodia), "Say No to Fake News Campaign," Ministry of Information, accessed August 2, 2025, <https://www.information.gov.kh/articles/85203>.
9. Cambodianess, "Rise in AI-Generated Scams Requires Quick Action," Cambodianess, accessed August 2, 2025, <https://cambodianess.com/article/rise-in-ai-generated-scams-requires-quick-action>.
10. Cambodia People's Party, "Media Outlets to Adhere to Journalism Charter," Cambodia People's Party, accessed August 2, 2025, <https://www.cpp.org.kh/en/details/410073>.
11. Ibid.
12. Cambodianess, "Media Outlets to Adhere to Journalism Charter," Cambodianess, accessed August 4, 2025, <https://cambodianess.com/article/media-outlets-to-adhere-to-journalism-charter>

13. *Khmer Times*, "Cambodia Readies for National AI Strategy," *Khmer Times*, accessed August 4, 2025, <https://www.khertimeskh.com/501673897/cambodia-readies-for-national-ai-strategy/>.
14. Rafique, Adnan, Furqan Rustam, Manideep Narra, Arif Mehmood, Ernesto Lee, and Imran Ashraf. 2022. "Comparative Analysis of Machine Learning Methods to Detect Fake News in an Urdu Language Corpus." *PeerJ Computer Science* 8 (June): e1004. <https://doi.org/10.7717/peerj-cs.1004>.
15. Mohsen, Fadi, Bedir Chaushi, Hamed Abdelhaq, Dimka Karastoyanova, and Kevin Wang. 2024. "Automated Detection of Misinformation: A Hybrid Approach for Fake News Detection." *Future Internet* 16 (10): 352. <https://doi.org/10.3390/fi16100352>.
16. Lakshmi Holla, and K S Kavitha. 2024. "An Improved Fake News Detection Model Using Hybrid Time Frequency-Inverse Document Frequency for Feature Extraction and AdaBoost Ensemble Model as a Classifier." *Journal of Advances in Information Technology* 15 (2): 202-11. <https://doi.org/10.12720/jait.15.2.202-211>.
17. Al-Tarawneh, Mutaz A. B, Omar Al-irr, Al-Maitah, Khaled S, Hassan Kanj, and Aly., 2024. "Enhancing Fake News Detection with Word Embedding: A Machine Learning and Deep Learning Approach." *Computers* 13 (9): 239. <https://doi.org/10.3390/computers13090239>.
18. Du, Jiangshu, Yingtong Dou, Congying Xia, Limeng Cui, Jing Ma, and Philip Yu. n.d. "Cross-Lingual COVID-19 Fake News Detection." Accessed June 27, 2024. <https://arxiv.org/pdf/2110.06495.pdf>.
19. Mohsen, Fadi, Bedir Chaushi, Hamed Abdelhaq, Dimka Karastoyanova, and Kevin Wang. 2024. "Automated Detection of Misinformation: A Hybrid Approach for Fake News Detection." *Future Internet* 16 (10): 352. <https://doi.org/10.3390/fi16100352>.
- 20.



Mr. Prohim Tam

Prohim Tam received his B.S. degree with the Department of Management of Information Systems from Paragon International University, Phnom Penh, Cambodia, in 2019. He completed his master's and Ph.D. degrees in the Department of Software Convergence at Soonchunhyang University in Asan, South Korea, in 2024. During his academic journey, he also began his IT intern job at the Cambodia Securities Exchange in 2017, worked as an application developer on integrated AI-based annotation techniques, abnormal detection systems, and policy recommendation models for the National Information Society Agency and the Korean Health Panel Study from 2020 to 2022, and has been actively contributing to R&D projects with Mobile Computing & Networking Lab since 2019 until now (under grants of IITP, NRF, and BK21). Currently, he lectures AI-cybersecurity at American University of Phnom Penh, Cambodia. He is passionate about research and has authored several articles published with IEEE, Springer, MDPI, Wiley, and other reputable peer-reviewed journals. He also actively contributes as a reviewer to the research community. His research interests include 5G/6G, privacy-preserving learning policies, applied AI in network services and management, and emerging network security in SDN/NFV.



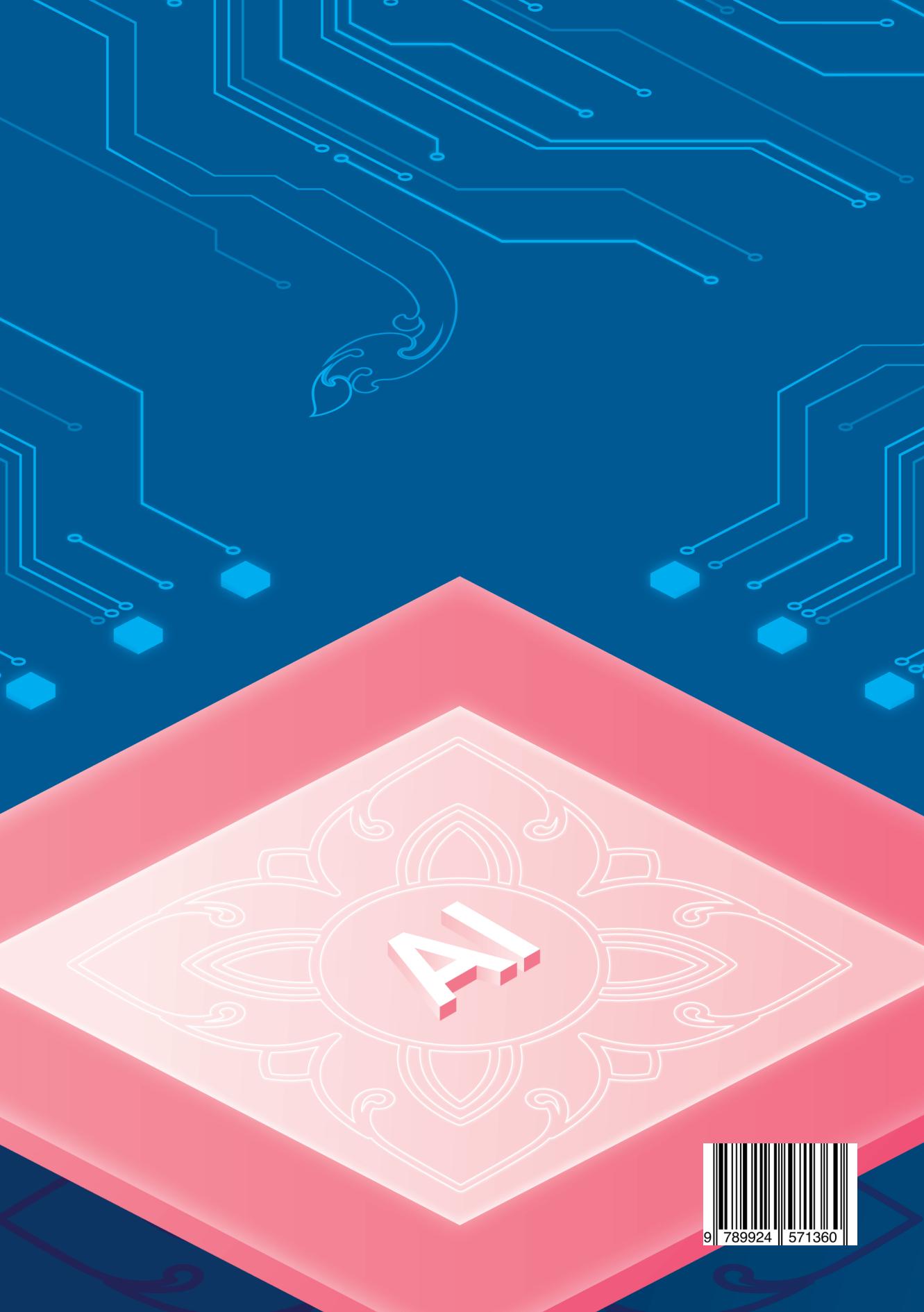
Mr. Riccardo Corrado

Riccardo Corrado, Ph.D. is an Associate Professor at CamEd Business School and serves as Vice Chairman of the Italian Business Chamber Cambodia (ItaCham), the Italian chapter within EuroCham Cambodia. He is also a Board Advisor for the Federation of Cambodian Students Abroad and AI Farm. Additionally, he has been working as a Consultant for the Telecommunication Regulator of Cambodia and the League of International Relations Affairs and is serving also as a Senior Research Fellow at the Asian Vision Institute.

Dr. Corrado holds a Ph.D. in Information Engineering, as well as a bachelor's degree in Electronics Engineering and a master's degree in Telecommunications Engineering from the University of Trieste, Italy. He also earned a Master's in Education from the University of Johannesburg, South Africa, and an MBA with a specialization in Business Analytics from Abertay University Dundee, UK.

His research focuses on the application of technology for sustainable development, regulatory frameworks in the telecommunications sector, and the digital transformation of public and private sectors, with a particular interest in the emerging space economy.

*This page is intentionally left blank.*



9 789924 571360