Discussion Paper Series 2025-26

DECOLONIAL STUDIES PROGRAM

Sana AI? Risk Mitigation in the Era of Big Tech and Surveillance Capitalism

Frances Cruz

DECOLONIAL STUDIES PROGRAM

Sana AI? Risk Mitigation in the Era of Big Tech and Surveillance Capitalism

Frances Cruz

UP CIDS Discussion Paper Series
is published by the

University of the Philippines
Center for Integrative and Development Studies
Lower Ground Floor, Ang Bahay ng Alumni
Magsaysay Avenue, University of the Philippines
Diliman, Quezon City 1101

cids.up.edu.ph

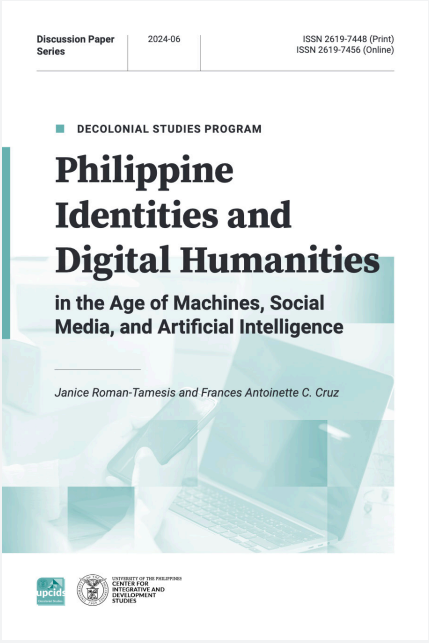
Telephone: (02) 8981-8500 loc. 4266 to 4268 / (02) 8426-0955
Email: cidspublications@up.edu.ph
Website: cids.up.edu.ph

ISSN 2619-7448 (Print)
ISSN 2619-7456 (Online)

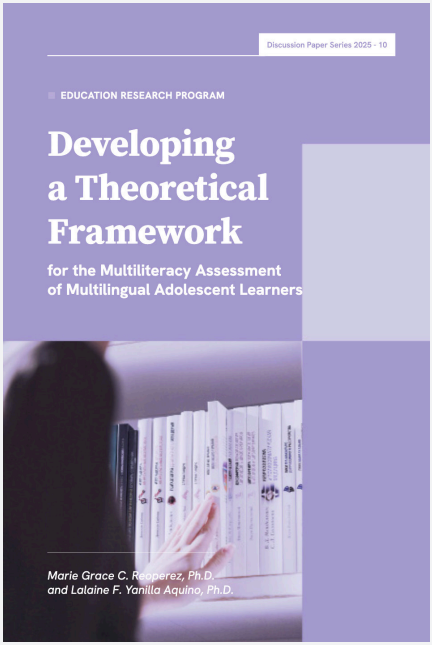
Cover image credit
"CCTV Cameras Guarded By Spikes" https://www.publicdomainpictures.net/en/view-image.php?image=256661&picture=cctv-cameras-guarded-by-spikes

Table of Contents

Download related policy papers for *free*



Philippine Identities and Digital Humanities in the Age of Machines, Social Media and Artificial Intelligence



Developing a Theoretical Framework: for the Multiliteracy Assessment of Multilingual Adolescent Learners

cids.up.edu.ph/collection-database

Get your policy papers published.

Download open-access articles.

The Philippine Journal of Public Policy: Interdisciplinary Development Perspectives (PJPP), the annual peer-reviewed journal of the UP Center for Integrative and Development Studies (UP CIDS), welcomes submissions in the form of full-length policy-oriented manuscripts, book reviews, essays, and commentaries. The PJPP provides a multidisciplinary forum for examining contemporary social, cultural, economic, and political issues in the Philippines and elsewh ere. Submissions are welcome year-around.

For more information, visit cids.up.edu.ph.
All issues/articles of the PJPP can be downloaded for free.

Get news and the latest publications.

Join our mailing list to get our publications delivered straight to your inbox! Also, you'll receive news of upcoming webinars and other updates.

bit.ly/signup_cids

We need your feedback.

Have our publications been useful? Tell us what you think.

bit.ly/dearcids

SANA AI? RISK MITIGATION IN THE ERA OF BIG TECH AND SURVEILLANCE CAPITALISM

Frances Cruz[1]

1 UP CIDS, Decolonial Studies Program/Department of European Languages, UP Diliman/ University of Antwerp/Ghent University. Contact: fccruz3@up.edu.ph

- The National AI Strategy Roadmap and the pending AI-related bills in the Philippine Congress demonstrate the urgency of a government response to the potential of AI. However, the pursuit of economic growth and the rush to provide training opportunities necessitates careful consideration of its misuse in terms of foreign surveillance, cybercrimes, and societal disharmony. This concern deserves urgent attention considering the importance of these in the National Security Policy.

- While capacity-building is important, readiness for the Fourth Industrial Revolution is not merely a matter of AI-related training but also of strengthening foundational skills, such as in reading, mathematics, and science, as well as critical literacy and critical digital literacy. The Philippines must improve its performance in these basic skills for a sustainable, skilled, and analytical future workforce.

- Besides policies geared toward the private sector, the government can work with civil society to understand the social impacts of AI and assist with the piloting of AI-related technology for marginalized groups, i.e., disaster management, weather prediction, crop yield prediction, and efficiency in water use in farms.

The following discussion paper is partially based on a roundtable discussion (RTD) on Artificial Intelligence (AI) and Decolonization that took place on Friday, 20 September 2024 at the University of the Philippines Center for Integrative and Development Studies (UP CIDS) organized by the Decolonial Studies Program. The roundtable, *Now You See IT*,[2] featured scholars and practitioners from the fields of law, journalism, data science, and governance, and was centered on the various ways that artificial intelligence intersects with the access and control of both technology companies and enabling states and governments have over private data, creative output, and decisions made with regard to individuals as a result of access to such data. Their input will be drawn on in this framing of pending AI legislation and related laws in the larger scope of AI risks and opportunities.

Increased accessibility to AI tools and the introduction of large language models (LLMs) such as OpenAI's GPT-4o, which the chatbot ChatGPT is based on, has brought about a significant shift in the ways humans seek information and interact with technology. As with many new technologies, the vast range of social, cultural, political, scientific, and economic dimensions of AI usage necessitate clear policies in public and private sectors, a roadmap to which was developed in 2021 by the Department of Science and Technology (DOST-PCIEERD 2021), with a National AI Strategy Roadmap 2.0 and the Center for AI Research being launched by the Department of Trade and Industry (DTI 2024). Some of the main foci of the first draft of the roadmap, as well as the many journal articles that have arisen in light of the country's adoption of AI-related policies, emphasize economic benefits, capacity-building, as well as different applications across fields such as farming, transportation, healthcare, disaster management, and education (Concepcion et al. 2019; Rosales et al. 2020; Estrellado and Miranda 2023; Gutierrez and Viacrusis 2023). This was echoed by the participants in the RTD. Dr. Maria Margarita Lavides of the Data Science for Public Policy Program at UP CIDS cited the relevance of AI for research and language research. Meanwhile, Professor Rachel Khan of the UP College of Mass Communication emphasized the advantages and disadvantages of any new technology, with the key determinant of either being the way technology is used. This question of how AI is used, by whom, and for whom is the crux

2 Found on https://cids.up.edu.ph/decolonization-decoloniality-ai-internet-roundtable/.

of assessing potential risks, particularly with regard to the connection of the use of AI with decolonization. Indeed, AI comes with both benefits and shortcomings.[3] Among its benefits, AI systems are faster and more efficient at consolidating information within and across databases, correcting for human errors in domains where precise measurements are needed, saving water in agricultural processes (Rosales et al. 2020), predicting crop yields (Kouadio et al. 2018), detecting cancer cells (Somers 2021), and developing new drugs for bacteria that have become resistant to antibiotics, such as in the case of Stanford Medicine's *SyntheMol* (Tompa 2024). On the other hand, AI has also been associated with cybersecurity threats, data privacy, environmental harms, intellectual property issues, job losses, lack of accountability, explainability, and transparency, as well as misinformation as noted by IBM (Caballar 2024), with the recent Association for Southeast Asian Nations' (ASEAN) *Guide on AI Governance and Ethics* noting potential risks include:

- Mistakes and anthromorphism;
- Factually inaccurate responses and disinformation;
- Deepfakes, impersonation, fraudulent and malicious activities;
- Infringement of intellectual property rights;
- Privacy and confidentiality;
- Propagation of embedded biases (ASEAN Secretariat 2024, 5)

While there are some differences in focus between the IBM list and the ASEAN guide, the latter focuses on the lack of accuracy, data and intellectual property protection, and cybercrime in the ASEAN guidelines. It is also nonbinding, acting solely as a guide for governments, as will be discussed below on the landscape of AI in the Philippines and the region below.

More recently in academia, the use of Chat GPT posed all sorts of challenges to educators, as ChatGPT can provide essay-length answers to prompts that could then be copied and presented falsely as one's individual creative work.

---

3 As one of the pioneers of AI, Geoffrey Hinton, told the MIT Technology Review in 2023: "I have suddenly switched my views on whether these things are going to be more intelligent than us. I think they're very close to it now and they will be much more intelligent than us in the future. [. . .] How do we survive that?" (Heaven 2023)

The irony of this is that AI itself requires copious amounts of input in order to produce what it does, human input based on other creative work. A large language model is fed by user-produced data, and through machine learning processes, whether neural networks, decision trees, regression, Bayesian approaches or otherwise, their usage in chatbots can produce answers that seem human-like. However, the ways these interfaces and Facebook can be abused by those in power unveil larger structural problems in the tech industry and its relationship with the Global South. In 2019, author Michael Kwet (2019) elaborated on the concept of digital colonialism, where he argues that tech giants (including but not limited to the "Big Five" comprising of Alphabet (Google), Amazon, Apple, Meta (Facebook, Instagram, and WhatsApp), and Microsoft[4] are influential in ecosystems in many parts of the world, while also extracting, and being able to extract, a host of personal data and information about individual users. These companies demonstrate a sheer value, with high market capitalization:, Microsoft at $3.399 trillion, Nvidia at $3.289 trillion, Apple at $2.993 trillion, Amazon at $2.173 trillion, Alphabet at $2.103trillion, and Meta at $1.618 trillion (Nasdaq for May 28, 2025, cited in Companies Market Cap 2025) Most of these companies have also begun developing their own AI tools that perform user-defined tasks, such as Alphabet (Gemini), Meta (Llama), OpenAI (ChatGPT), and Microsoft (Copilot).

Some of the noticeable societal impacts of these in the region besides data privacy concerns are misinformation, which can range from phenomena such as AI hallucinations (for instance, the tendency for chatbots to cite fake sources in their output), the way that the architecture of social media already customizes feeds to reflect posts that would cause users to engage with content that aligns with their own interests (Han 2018), as well as the environmental harm of AI tasks. One study estimates that the training of a large transformer with neural architecture search emits around 626,155 pounds of carbon dioxide, which is worth more than 56 years of the carbon dioxide emissions of an average human individual (Strubell, Ganesh, and McCallum. 2019). Considering the disproportionate impacts of climate change on developing (alternatively, Global South or low- and middle-income countries) (Adom

4 This group may also include Nvidia, with its market capitalization of $3.395 trillion as of writing (Companies Market Cap 2024) as part of the "Big Six."

2024), there is all the more reason for critical perspectives to emerge from such contexts as will be discussed below.

In line with this, the antecedents of AI risks largely have to do with the context that provides power to Big Tech companies in the region, and by extension, globally. One of the ubiquitous examples of the scope of this power is the way Free Facebook, or Free Basics, later called Discover (Tobin 2021), was promoted in developing countries. Free Basics was a program of Facebook, and later, Meta, that was piloted throughout many countries in the developing world, to address concerns about the "digital divide" which limited the accessibility of the internet to millions of people. Without needing a data plan, users of Free Basics could gain access to Facebook (FB) and a small collection of websites (basics) through their phones, and both Free Basics and the implementation of Discover included countries in Southeast Asia such as Thailand, Indonesia, and the Philippines (Tobin 2021; Roth 2022). The weight of this increased accessibility cannot be understated in the Philippines, where there were are an estimated 83.85 million Facebook users in a population of 111.8 million in 2022 (Kemp 2022). Additionally, the Philippines has one of the highest social media usage rates in the world, second only to South Africa (Baclig 2022). The number of social media users was so much so that in an interview with Time magazine, Maria Ressa, former CNN journalist and Rappler's CEO once said of Facebook, "In my country, Facebook essentially is the Internet, thanks to subsidies from telecommunications companies that let people avoid data charges while on the site" (Ressa 2019). It was in the context of the Duterte administration that this interview was conducted, amidst the growing political instrumentalization of social media (Curato 2021). During the course of the presidency, trolls were documented to have been linked to campaigns in support of the administration, by the admission of the social media manager of the Duterte campaign, Nic Gabunada (Cabañes and Cornelio 2017; Bernado 2018; Curato 2021). For Cabañes and Cornelio (2017, 233–35), a troll employs aggressive and abusive posts or messages online to forward a political agenda and may refer to an individual speaking about their own political convictions or paid individuals with an individual account or multiple accounts. For many, this was an alarming development as it appeared to be yet another medium where hegemonic or even traditional political processes could be performed and amplified through programs such as Free Basics, wider internet reach, smartphone availability, and so on.

The spread of such discourses is, of course, not directly related to AI, rather they serve as the social backdrop for a scandal that used the social media architecture and the personal data of users together with AI for political purposes. Netflix's documentary The Great Hack (2019) recounts the history of the Cambridge Analytica company, which had been the subject of various news reports alleging that it had been harvesting personal data on Facebook without consent. These reports were followed by a complaint filed at the Information Commissioner's Office in 2017 by Professor David Carroll, who sought to reclaim his personal data from the company, as well as investigative journalist Carole Cadwalladr's interview of the whistle-blower Christopher Wylie, who elaborated on Cambridge Analytica's illicit gathering of social media data once reported in the media.

At the heart of the Cambridge Analytica scandal was the use and instrumentalization of personal data by private companies or third parties for political purposes. Theoretically, the collection of data was based on a theoretical assumption detailed in Kosinski, Stillwell, and Graepel (2013) that linked certain personal characteristics to digital information. Alexander Nix of Cambridge Analytica contracted Aleksandr Kogan, a colleague of Kosinski at the University of Cambridge, Aleksandr Kogan, was then contracted by Alexander Nix of Cambridge Analytica to operationalize Kosinski, Stillwell, and Graepel's (2013) theoretical assumptions through the development of a quiz app, "thisisyourdigitallife," which gathered the personal data of those who took the quiz, as well as their friends (Cadwalladr and Graham-Harrison 2018). The data gathered was then used to identify users whose personal information could act as predictors of behavior, namely the likelihood of being persuaded by political posts. Rather than select users who were already convinced of their political choices, the company was able to target those users in "swing" areas in election campaigns where just a few votes could make a difference.

The decision of Professor Carroll to invoke the rights to his own data recalls the butterfly effect, in which minute individual decisions can potentially have enormous reverberations in the unfolding of events. Both he and journalists such as Carole Cadwalladr, Harry Davies, Hannes Grasseger, and Mikael Krogerus, whistleblowers such as Brittany Kaiser and Christopher Wylie, and many others contributed to raising public awareness about the dangers of illicit data-gathering and restrictions of access in Facebook's Graph API. However, Cambridge Analytica eventually led to one firm and the malicious

use of personal data is unlikely to stop after this scandal. Returning to the matter of context, the proliferation of AI used in contexts of influencing populations and elections occurs in a greater context of continuing global core-periphery relations where the global core extracts labor and resources from the periphery. This is certainly not an old idea, and theorists such as Wallerstein (1976) have written on these dynamics extensively. This context is all the more important to grasp in the Philippines and developing countries more generally, as the country finds itself needing to adapt to rapidly changing technologies while managing the challenges brought about by such. The following sections thus focus on two aspects of AI: the context of technological dependency and surveillance with regard to decision-making, as well as discourse and misinformation.

THE AUTOMATED I/EYE

The first concerns brought up in this discussion are related: dependency and surveillance. The domains of Microsoft, Amazon, Meta, and Google are global in scope, with a report released by Human Rights Watch detailing their reach with almost three billion people using Facebook, WhatsApp, or Instagram daily, 90 percent of the world using Google Search, and Google's Android is the operating system of about 75 percent of the world's smartphones (Brown 2020). In addition, Amazon Web Services, Microsoft Azure, and Google Cloud Platform are the three most dominant providers of cloud computing services as of writing. This dominance means that many quotidian functions of everyday (digital) life are dependent on large private technology companies, from emailing to streaming to shopping to the storage of personal, educational, and professional information, and so on. With relevance to the developing world, Kwet (2019) shows how under the aegis of bridging the digital gap, the South African government negotiated with various corporations and the educational sector under NDAs for the digitalization of classrooms for Operation Phakisa Education. The secrecy with which the talks proceeded hindered critical inquiry as to the risks and impacts involved in the large-scale implementation of technological services, many of which, if not all, are largely centered in the United States or the Global North due to structural inequalities (Yu, Rosenberg, and Gupta 2023). Kwet (2017) further noted the historical ways in which surveillance technology was used to support and enable South African apartheid, as well as ways in which new technologies can have a chilling effect on free speech as well as teacher autonomy through the misuse of monitoring

capabilities. While the secrecy involved in the Operation Phakisa discussions on educational opportunities and services offered by large tech companies makes the involvement of AI in such processes a matter of speculation, the ability of these services to store grades, curricula, demographic information that can be gleaned from assignments such as personal essays, and so on, cannot only be made available for AI training based on the decisions of the parent company. Their potential for misuse has been exemplified in the automated grading scandal for A-levels in the United Kingdom due to the cancellation of in-person exams during the COVID-19 pandemic. Using three inputs, namely the grade distribution of schools from 2017 to 2019, teacher evaluations of student ranks per subject, and previous exam results of the students by subject, an algorithm generated an estimated A-level grade that disproportionately affected students from low-income backgrounds and was roundly criticized for questionable statistical assumptions and low algorithm accuracy (Kolkman 2020). The State and other forms of private security might potentially misuse these often poorly understood processes, and this has already happened. Crawford (2021) detailed how automatized processes have often largely replicated human biases. This observation is also exemplified in the Global South context in Rahman (2023), who remarks that front-office employees or clerks do not often question or override automated decisions, and how it takes away the human agency exercised in gray areas or when faced with technical errors. In other words, algorithms employed for decision-making tend to generalize prejudices that are seen repeatedly in the datasets on which algorithms are trained, and are therefore, without any oversight or training/parameter adjustment, agents of mass bias replication against people fitting certain profiles. On a more ontological level, certain forms of databases tend to assume identities as fixed, depending on the nature and flexibility of data gathering (Rahman 2023). This characteristic can be particularly concerning. Where the prediction of human behavior is linked to decisions about humans, it is unknown if an algorithm can properly predict or conceptualize human agency and the capacity to change, and our own theoretical understanding of such is limited. The utilization of such algorithms has real-life effects once they are uncritically adopted by firms, particularly those that do not subject the algorithm's code to checks and balances often present in democratic governments around the world. The datasets these were trained on are also rarely critically assessed for biases, omissions (both of entries and variables) and data quality.

MASS SURVEILLANCE AND DEPENDENCY

A related aspect of the dependency on tech companies is their relationship with mass surveillance. In the section above, I discussed how such systems can and have already been exploited for the purchase of personal data, with data becoming an increasingly sought-after and expensive resource (*The Economist* 2017). For the domain of surveillance, the Philippines is an apt starting point, as it was in the Philippines that the United States was known to have experimented with applying surveillance technologies over a large population, beginning as early as 1903. These came in the form of data management techniques such as fingerprinting, photographs, and indexing, and a grid made up of "5,355 miles of landlines and 1,615 miles of undersea cable" (McCoy 2009, 108). McCoy goes on to argue that the Manila Metropolitan Police, the Philippine Constabulary, and the Division of Military Intelligence were then used for monitoring, censoring, infiltrating, and surveilling civil society to ensure compliance with the American regime (110). While these actions were largely carried out in the capital and other areas that belonged under the United States' "civil regime," then the areas of the current territory of the Philippines that had not previously come under Spanish control, namely the Cordilleras and much of BARMM, were subject to an even more rigid and controlled "military regime" (Abinales 2010). The personal data that is available and published through various tech companies has not only been abused for commercial purposes as shown above but has been used to train various algorithms for classification.

When placed in the context of surveillance, classification algorithms can automate processes based on training data that were formally largely in the hands of humans at a larger scale. Regardless of whether machine learning is supervised or unsupervised, classification algorithms tend to reflect biases in human classification, as can be seen in the work of Bolukbasi et al. (2016), aptly titled *Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings*. The title itself is a reference to the capabilities of word embeddings, which are representations of words in vector space that allow for the identification of words that are used in similar contexts. Such representations allow for equations that show different types of relationships between words and their contexts of use, with the title above giving an example of one such result that is heavily influenced by the ideas and contexts found in the training dataset.

Once this type of dataset content is fed into large language models, existing social biases are amplified or simply left unquestioned if those who implement or program the algorithms do not have the knowledge or capacity to make necessary adjustments—i.e., the technology, left unchecked or in the hands of those without sufficient background knowledge, does little to progress the status quo, and may even automate discriminatory practices that are prohibited in law without the intention of the implementer. Some of the more critical effects of this type of classification can be seen when automated classification is applied to a vulnerable minority group. This was seen in the case of the Rohingya refugees in Bangladesh, who went on strike in 2018 to demand their own ethnic name be placed on their Smart Card (as opposed to belonging to either Myanmar or Bangladesh), and for the authorities in Bangladesh not to share their collected biodata with the Myanmar government in light of the possible threats to their lives (Rahman 2023, 29). Yet another alarming development in this sector concerns autonomous weapons systems have been the subject of the *Campaign to Stop Killer Robots* since 2013, as these run "without meaningful human control" (Stauffer 2020). Even while AI is the specific target of such critiques, the context in which AI can amplify specific biases is crucial. In situating the context of AI, Shoshana Zuboff writes of surveillance capitalism, which

> unilaterally claims human experience as free raw material for translation into behavioral data. Although some of these data are applied to product or service improvement, the rest are declared as a proprietary behavioral surplus, fed into advanced manufacturing processes known as "machine intelligence," and fabricated into prediction products that anticipate what you will do now, soon, and later. Finally, these prediction products are traded in a new kind of marketplace for behavioral predictions that I call behavioral futures markets. (2019, 4)

AI is not only related to predictions of behavior for marketing or societal purposes but has sometimes been used in matters of life or death. As of writing, news reports have uncovered in the light of the assassination of the CEO of medical insurance company United Healthcare (UHC) in New York how the company was sued for allegedly using algorithms with a 90 percent error rate to deny medical claims even with doctor's approval (Al-Sibai 2024). In the domains of law and order and security, a trial that employed AI to identify the faces of previously unknown suspects was conducted by the London

Metropolitan Police in 2018 and was found to be 98 percent inaccurate, with concerns raised about discrimination along race, sexual orientation, and age (Sharman 2018). Furthermore, in 2021, Google and Amazon signed a contract known as "Project Nimbus" that provided cloud computing infrastructure and AI amongst other computing services to the Israeli military (Al Jazeera Staff 2024). When combined with assumptions linking human attitudes to online behavior such as those articulated by the Kosinski study (2013), there are disturbing possibilities when implementing automated systems for policing or military purposes for treaties, particularly with regard to upholding the provisions of the Rome Statute, and in more local contexts, the legal principle of the presumption of innocence pending due process. As shall be discussed below, the dangers of using AI for procedures that will ultimately make life-changing decisions for individuals and communities is one of the components of laws intended to regulate AI.

DISCOURSE AND AI-DENTITY

This second section deals with discourse and its relationship to AI. Although there are many definitions of discourse, here we will refer to discourse as a way of organizing knowledge to structure realities (Foucault 1972), or as a form of social practice (Fairclough 1992, 1995).

Critical studies of discourse often critically examine the nexus between various texts, power, and ideology, showing how juxtapositions of meanings can accrue and disseminate throughout various forms of media, forming the basis of what can be "said" and what is marginalized (Wodak and Meyer 2016). The irresponsible use of AI in politics, as noted above and also by Geffrey Hinton (Heaven 2023) and *The Great Hack* (Noujaim and Amer 2019), has the potential to produce a semblance of consensus or majority opinion when there is none. This could also spread disinformation and foster greater societal, cultural, and political polarization (Bradshaw and Howard 2017; Cabañes and Cornelio 2017; Curato 2021; Lee and Theokary 2021). This concern was similarly the focus of Ferdinand Sanchez's contribution to the RTD. Not only do predictive algorithms that enable divisiveness rely on a type of surveillance or automated tracking of data subjects, as discussed above, but also classifier algorithms that determine suitable content within the purview of what a particular company or state deems as such have also blurred lines between

what constitutes "appropriateness."[5] In 2020, a Human Rights Watch report noted that videos of human rights abuses in Syria and Yemen that were posted online were censored, hindering efforts at documentation procedures in aid of investigation and attaining justice for victims (Brown 2020). If large language models, or algorithms in general, are responsible for curating the selection of accounts or posts that appear in one's social media feed, in other words, acting as an information gatekeeper (Shoemaker and Han 2020), then they increase not just the likelihood for generating sameness but also reduce the likelihood of an encounter with the "other" that only enables political polarization (Brown 2020; Han 2018). In a country such as the Philippines, where political opinion is often split between reformist and populist approaches (Thompson 2010) and the "civil" and "mass" spheres (Kusaka 2017), polarization only decreases the opportunities for "zones of contact" or community spaces that aim towards more productive activities and discussions to address local concerns (Kusaka 2017).

THE POLICY LANDSCAPE FOR AI

In terms of the policy landscape for AI, the regional context saw non-binding guidelines for AI being developed by ASEAN, as mentioned above and released on 24 February 2024. Though the guidelines draw from other documents focusing on AI on a multilateral scale such as *UNESCO's Recommendation on the Ethics of Artificial Intelligence and the EU's Guidelines for Trustworthy AI,* they do not go so far as to set binding legal precedents for countries in ASEAN, in alignment with the "ASEAN Way' of non-legalistic, harmonious conduct in multilateral affairs (Goh 2003). The guidelines, shy of the world's first comprehensive AI law, the EU AI Act, which came into force on 1 August 2024, provide legal and regulatory provisions for the use of AI within the European Union (EU). The act proposes a five-tier risk categorization system for AIs with regard to their potential to violate fundamental human rights and freedoms. The five-tier system includes (a) unacceptable risk, (b) high risk, (c) general-purpose AI, (d) limited risk, and (e) minimal risk. Notable exemptions to this include AI used for military or research purposes, as discussed above.

5 Arguably, location tracking may be turned off for particular applications, yet this inevitably requires greater awareness on the part of the user of the consequences of tracking and the often cumbersome procedures for turning off default or previously enabled tracking.

The reaction to the EU's new AI law, while providing safeguards for the irresponsible implementation of AI and potential deleterious impacts on human rights, was critiqued for its inability to address the "root causes" of the existing issues with AI, such as "its role in increasing and entrenching the extreme power a few dominant tech firms already have in our personal lives, our economies, and our democracies," as Max von Thun, Europe director for the Open Market Institutes said in an interview with *Euronews* (Davies 2024). As argued above, while this direct regulatory approach is not envisioned in the guidelines issued by ASEAN, the guidelines assert that protect and strengthen the rights of individuals pre-established in other laws: "Given the profound impact that AI potentially brings to organizations and individuals in ASEAN, it is important that the decisions made by AI are aligned with national and corporate values, as well as broader ethical and social norms." (ASEAN Secretariat 2024, 7).

Though regulation has not been the primary focus of ASEAN, as of writing, there are several bills in the Philippines pending in Congress about AI. These include House Bill (HB) 9448, An Act Regulating The Use of Artificial Intelligence and Automation Systems in the Labor Industry and for Other Purposes (Juan Carlos Atayde); HB 7396, An Act Promoting the Development and Regulation of Artificial Intelligence in the Philippines (Robert Ace Barbers); HB 7913, An Act Establishing a Regulatory Framework for a Robust, Reliable, and Trustworthy Development, Application, and Use of Artificial Intelligence (AI) Systems, Creating the Philippine Council on Artificial Intelligence, Delineating the Roles of Various Government Agencies, Defining and Penalizing Certain Prohibited Acts (Keith Micah Tan); and HB 7983, An Act Providing a National Strategy for the Development of Artificial Intelligence, Creating for the Purpose the National Center for Artificial Intelligence Research, and Appropriating Funds Therefor (Keith Micah Tan). All these bills are concerned with regulatory mechanisms in some way. A challenge that these laws present has come up with regard to the definition of AI, which is, as of writing, not formally defined in Philippine law. Atty. Emerson Bañez (2024), one of the roundtable discussants, variously characterizes the definitions of AI based on the pending bills as 'sophistication of the programming' or the methods involved in AI, such as machine learning. Yet Bañez also remarks that AI ought not to be the domain of private companies, rather the onus is on the public sector to ensure technological accessibility. This echoes the assertion of Katharina Zügel, policy manager at the Forum on Information

and Democracy, who argued in response to the EU AI Act that AI ought to be a public good (Davies 2024).

While the concern with balancing the potential risks of AI with potential benefits is a welcome step in legislation, which has historically and on a global level been relatively slow to respond to developments in technology, the AI-specific bills that are still up for discussion ought not to obscure existing laws and frameworks that are relevant to some of the issues mentioned above (Divina and Ipac, 2024). These include the implementing rules and regulations of the Data Privacy Act of 2012, such as the data subject's rights in accessing "information on automated processes where the data will, or is likely to, be made as the sole basis for any decision that significantly affects, or will affect, the data subject" (Section 34.c.2), and one's right of objection to the processing of their personal data (Section 34.b) (National Privacy Commission 2016). Enabling legislation for training in AI also exists, albeit is focused on developing skills and competitiveness for labor rather than exercising critical awareness of potentially harmful societal impacts of technology in the form of Republic Act 11899, the Second Congressional Commission on Education Act II, which includes among its provisions the enhancement of the skills and competitiveness of the Philippine workforce in human and digital technology and innovations (Divina and Ipac 2024). The Philippines has further participated in multilateral initiatives with regard to the use of AI. The country adopted the UNESCO recommendation on AI ethics in 2021 and the Bletchley Declaration in 2023 while advocating for international guidelines in terms of the ethical use of such weaponry. The Philippines is currently drafting a protocol and roadmap with other countries to "outlaw fully autonomous weapon systems under the Convention on Conventional Weapons" (Ordoñez 2023). What these provisions have in common is their focus on human rights, guiding AI ethics and the general principles behind regulatory legislation.

A potential connection between AI laws and other policies in the Philippines that requires attention is the context of security. The administrations of the Philippines since Benigno Aquino III have adopted a National Security Policy (NSP) that draws on the ASEAN-developed notions of comprehensive security, which considers security as a multidimensional process where socioeconomic conditions are inextricable from physical or traditional security (Caballero-Anthony and Emmers 2006). Thus, among the consistent NSP pillars are the ideas of "cultural cohesiveness" alongside "moral-(spiritual) consensus" and

a continuing focus on cybercrime and the environment (National Security Council 2011, 2017, 2023). As national security policy highlights these as priority efforts, their intersection with the deleterious effect of AI on the environment, risks of data privacy, and sociocultural polarization are all the more crucial and may also require "legal innovation" (Bañez 2024) emphasizing proactive critical and sustainable measures instead of merely reactive resiliency measures. The issue of polarization is reflected in the current NSP, which states:

> The Government shall take heed of the current strategic environment where advancements in online and digital learning technologies are engendering cultural divides and social conflicts. Today's colonization does not just occur through inter-governmental and global action but also on a societal level through individuals' thoughts, actions and interactions. Online platforms shall be utilized to accelerate the acquisition of holistic knowledge and experience and to understand what is meaningful and valuable. (National Security Council 2023, 29)

What can be gleaned from these relationships is that AI is not a one-policy issue, rather, requires a more comprehensive approach to mitigate risks while promoting development. As highlighted above, what can be constituted as an appropriate policy response amidst technological inequalities is the recognition and incorporation of local contexts, cultures, visions, values, and capabilities in its framing.

PLACE AND AI

One of the more recent initiatives in the Philippines is the ePLDT Pilipinas Cloud, which was created based on the concept of digital sovereignty that was invigorated in discussions about the United States National Security Agency's PRISM program, which was alleged by reports by Edward Snowden to have accessed personal data of individuals from outside the United States through American internet companies (Padilla 2014). Technological countermeasures under the framework of digital sovereignty were then proposed and discussed by states outside the United States to ensure additional protection for their citizen's data, or for the data shared in or through large tech companies. The possession of a sovereign cloud allows for the storage of sensitive data of the citizens of a particular country and is subject to local regulations. This is

arguably a crucial development in ensuring adequate protection for Philippine data subjects, even when representing just one of the myriad safeguards to be put in place for AI-related surveillance, data collection, and potential misuse. However, the matter of protecting Philippine user data from potential misuse by foreign governments, while implied in the Data Privacy Act's Implementing Rules and Regulations, does not go so far as to establish separate processing facilities for social media or tech companies with popular applications within the Philippines, such as is being discussed for Project Texas for American users of TikTok (TikTok 2024).

The ability of states or large regional blocks such as the EU only makes states and regional blocks of the developing world more vulnerable to abuses of technological-driven surveillance and decision-making by firms in the Global North. This is particularly evident in the way that the concept of the "digital divide" has been discussed multilaterally and as a form of development, in that it presumes that states in the Global South quickly have to play "catch-up" with the Global North in a quasilinear path of development (Graham 2011). Though there should certainly be initiatives to remain up-to-date with technological education, in the rush to do so, considerable time and care must be taken to acknowledge the agency of state and nonstate actors in the Global South to recognize alternatives to technological dependencies while assessing the needs of their own societies holistically. For instance, stakeholders in the Global South could account for the massive costs for the continuous upgrading, maintenance, and safeguarding of hardware such as PCs, the practical realities of rushed capacity-building, as well as the unsustainable amount of garbage generated by discarded devices which are known to generate and compound environmental problems.

A German term, *Torschlusspanik*, which can be thought of as the fear of time running out to perform a specific action, can arguably be related to decisions made under immense pressure to prepare society for AI readiness, or digital literacy, or earlier, ICT (Cueto et al. 2023). Though there are issues involving AI that are certainly of great urgency such as the matters of data privacy, intellectual property, ethics, and AI's use in matters of cybercrime as noted above (see for instance Arcilla et al. 2023), other aspects of new technologies require greater long-term planning and strategy to be implemented successfully, while properly situating the incorporation of technologies into existing laws, citizen's rights, local contexts, customs and capabilities, and

comprehensive visions of a "good society" (living standards, environment, and so on). If there are indeed temporal, spatial, and contextual inequalities between the Global North and Global South about matters such as AI, then there is also a form of shared learning where the Global South can adjust policies accordingly based on early adopters within the Global South (in high-income countries), as well as the experiences of the "Global South within the Global North," or marginalized communities within high-income countries. Indeed, the distinction between the Global South and Global North often obscures other transnational "digital divides," as well as other dissimilarities of context within the Global North and South (Graham 2011). Although structural issues persist in pursuing some of the more expensive solutions to protect data privacy and potentially, national security, such as the plans for Project Texas (TikTok 2024), individual and community-level discussions and initiatives are equally crucial in the face of a sector experiencing such rapid development as tech.

Apart from legal innovations that consider technological possibilities for achieving less dependency on large companies often situated outside of the legal jurisdiction of the country, existing and future policies ought to take into consideration aspects of their own implementation where automated dataset-driven algorithmic decision-making may have relevance and tie this to preexisting laws for data privacy. Yet another area of potential policy and practice in terms of AI is education. Republic Act No. 11899 (EDCOM 2) prioritizes the "adoption of digital transformation in education, and the use of science, technology and innovation through the promotion of digital literacy, and development of critical thinking, problem-solving and other related core competencies at par with global standards" (2022). The implementation of critical thinking and digital literacy, however, remains challenging in the Philippines. A report cited that less than half of the Philippine population have at least one of the six ICT skills of using basic arithmetic formulas in a spreadsheet; copying, pasting, and moving information in a document; sending emails with attachments; creating presentations with software; transferring files between devices; and installing and configuring software (Albert 2021). Besides the challenges in digital literacy, there are also lingering challenges with introducing critical literacy. As both critical digital literacy and critical literacy are related and presuppose both functional literacy and functional digital literacy, EDCOM 2's implementation requires strengthening the foundations of digital literacy even as it calls for a form of critical digital

literacy. The ways that AI can aid in this, for instance, are hinted at in the country's AI roadmap (DOST-PCIEERD 2021), as well as by Dr. Lavides during the RTD, who pointed out the utility of AI for research in Philippine languages and concepts: developing LLMs for chatbots or other forms of culturally-oriented generative AI in the Philippines, for example, requires the collection of datasets featuring Philippine languages and other cultural objects.

More importantly, fostering greater performance in the indices for digital literacy is dependent upon laws and policies that on the surface level seem to have little to do with AI or digital technology. Fajardo (2016), for instance, reports that teachers often shared that much was expected of them to teach critical literacy without providing adequate time off for the necessary teacher training or for the curation of culturally appropriate materials for rapidly changing educational needs. Issues with differing levels of digital literacy that are materially driven also bring about questions about the appropriate time or subjects in already-saturated curricula for introducing the concepts of critical literacy and critical digital literacy, not to mention technical upskilling if this is envisioned for primary or secondary education. Regarding the point earlier made about transnational similarities and dissimilarities of context that go beyond spatial concepts of the Global North or Global South, further compounding the issue of digital literacy are existing structural inequalities within the Philippines that are mirrored in other states. These concerns include inequalities between urban and rural connectivity, class-related aspects of access to PCs or laptops, as well as prohibitive costs and accessibility of utilities and devices for the training of students nationwide. In a meta-study on the implementation of ed-tech in the Global South, Cueto et al. (2023) found that in the rush to provide equipment, countries often did not have long-term plans or outcomes for such, as with the *Torschlusspanik* described above. Secondly, the authors recommended that educational policy ought to be on educational goals as opposed to merely technological goals, as well as the safety and privacy of students (Cueto et al. 2023, 51).

While the above passages have noted some of the observations and developments within AI-related policymaking, community-level and individual initiatives help cultivate skills and critical literacies. As Acerbi (2020, 213) writes, in terms of one's own engagement in online discourse and sharing one's personal information, "it is up to us." Community organizing and awareness bring about checks and balances necessary to create an alternative

set of practices to top-down initiatives, which may be hastily pursued to play catch-up without regard for appropriateness or viability in Global South contexts and in consideration of immense inequalities in infrastructure, policy, and funding. In suggesting pushbacks against some of the more deleterious risks of the domain of Big Tech companies and AI, Kwet (2019) notes the alternatives provided by FreedomBox, which allows one to run a personal server with its own privacy protection. Institutions, communities, and initiatives that also support and encourage free and open software are similar ways in which monopolies can be countered, and fewer opportunities are created for potential breaches in personal and creative data collection, processing, and training for AI (Kwet 2019). All in all, with rapid developments in technology, the pressure for quick policy responses, particularly when faced with industry demands, ought to be balanced with context-appropriate plans and strategies at the state, nonstate, and individual levels. These should be guided by local knowledge and research, drawing from successful strategies implemented in similar sociopolitical, cultural, and economic contexts that center and reinforce both established educational and development goals, as well as the rights of data subjects. Many of the implementation issues that arise with AI and the larger context of the digital divide are related to the successful implementation of other policies, such as those that have to do with teacher training and capacity-building; communication infrastructure; recycling, sustainability, and waste management policies; environmental policies; accessible, affordable utilities; and education. The proposed plans of action are:

1. Assure long-term policies to increase performance in basic reading, science, and numerical skills to create a sustainable pool of adequately critical and skilled talent for the "fourth industrial revolution." The Philippines performs relatively poorly in reading, science and mathematics in comparison to the world, and the region as evidenced by PISA 2022 (OECD 2023) (ranking 77 out of 81 countries) and the SEA-PLM 2019 Main Regional Report (UNICEF and SEAMEO 2020), while also having poor digital literacy skills (Albert 2021). All of these skills are necessary for citizens who can critically appraise and are technically capable of harnessing AI towards progressive societal goals.

2. Ensure equitable access to adequate telecommunications infrastructure and reduce costs of access to the internet. Around 47 percent of public

school students in rural areas do not have regular internet access due to high prices and unstable internet connection (Dalofin 2022). In particular the of cell towers through DICT's framework of Connect, Harness, Innovate, and Protect (CHIP) needs to expedite their construction where possible (i.e. through red-tape reduction) (Dalofin 2022).

3. Create mechanisms implemented at the local government level for proper e-waste collection and increase the number of treatment, storage, and disposal facilities in anticipation of greater e-waste generation.

4. As suggested in Arcilla et al. (2023, 325), ensure proper regulatory mechanisms for ethical AI to uphold public good, and to promote monitoring boards that comprise input and consultation from "multiple stakeholders, such as developers, users, regulators, and civil society organizations."

5. Identify and prioritize training, funds, and infrastructure development towards community-level grassroots initiatives where AI can be implemented for the betterment of the lives of the marginalized. This can proceed based on pilot projects for post-hoc assessment, improvement, and possible scale-ups, i.e., for water use efficiency, crop yield, disaster warning systems, weather prediction, translation for public administration and legal services, healthcare assistance, etc.

6. Fund research and/or monitoring groups that collaborate with or involve the appropriate government agencies and the stakeholders mentioned in point four above to periodically assess educational and training content alongside employment opportunities with the rise of AI. Such incentives not only can propose measures to reduce and mitigate the potential loss of employment and instability within the labor sector but can strengthen civil society and labor groups to adequately address the misuse of AI in public and private sectors.

These proposals are, of course, just a few in the growing literature on AI policies within the Philippines. As many of such proposals are geared towards economic benefits and immediate educational responses through capacity-building (i.e. through public-private partnerships), "readiness for AI" is, in the long-term, managed more holistically and comprehensively if the

public sector does not lose sight of democratic values and strengthening its educational and societal foundations.

REFERENCES

Abinales, Patricio. 2010. *Orthodoxy and History in the Muslim-Mindanao Narrative*. Ateneo de Manila University Press.

Acerbi, Alberto. 2020. *Cultural Evolution in the Digital Age*. Oxford University Press. https://doi.org/10.1093/oso/9780198835943.001.0001.

Adom, Philip Kofi. 2024. "The Socioeconomic Impact of Climate Change in Developing Countries in the Next Decades: A Review." Working Paper 681. Center for Global Development, February 2024. https://www.cgdev.org/sites/default/files/socioeconomic-impact-climate-change-developing-countries-next-decades.pdf.

Al Jazeera Staff. 2024, April 23). "What Is Project Nimbus, and Why Are Google Workers Protesting Israel Deal?" Al Jazeera, 23 April 2024. https://www.aljazeera.com/news/2024/4/23/what-is-project-nimbus-and-why-are-google-workers-protesting-israel-deal.

Albert, Jose Ramon G. 2021, October). "Why Literacy Measurement Deserves Rethinking." *PIDS Policy Notes* No. 2021-10, October 2021. Philippine Institute for Development Studies. https://pidswebs.pids.gov.ph/CDN/PUBLICATIONS/pidspn2110.pdf

Al-Sibai, Noor. 2024. "Murdered Insurance CEO Had Deployed an AI to Automatically Deny Benefits for Sick People." Yahoo! News, 5 December 2024. https://www.yahoo.com/news/murdered-insurance-ceo-had-deployed-175638581.html

Arcilla Jr., Angelis O., Adrian Kim V. Espallardo, Clio Andrei J. Gomez, Edrian Miles P. Viado, Vince Jeremy T. Ladion, and Ruben Andrei T. Naanep. 2023. "Ethics in AI Governance: Comparative Analysis, Implication, and Policy Recommendations for the Philippines." *Proceedings of the 2023 27th International Computer Science and Engineering Conference (ICSEC)*. 27th International Computer Science and Engineering Conference (ICSEC), Koh Samui, 14–15 September 2023. https://doi.org/10.1109/ICSEC59635.2023.10329756.

ASEAN Secretariat. 2024. *ASEAN Guide on AI Governance and Ethics*. Association of Southeast Asian Nations. https://asean.org/book/asean-guide-on-ai-governance-and-ethics/.

Baclig, Cristina Eloisa. 2022. "Social Media, Internet Craze Keep PH on Top 2 of World List." *Philippine Daily Inquirer*, 29 April 2022. https://newsinfo.inquirer.net/1589845/social-media-internet-craze-keep-ph-on-top-2-of-world-list

Bañez, Emerson. 2024. "Notes on Pending AI Legislation." *Rappler*, 24 January 2024. https://www.rappler.com/voices/imho/opinion-notes-on-pending-artificial-intelligence-legislation/

Bernado, B. 2018. “Philippine Media Under Attack.” In J. del Rosario-Malonzo (Ed.), In *State Terror & Tyranny in the Philippines: US-Backed Oplan Kapayapaan & Duterte’s Attacks on People’s Rights*, edited by Jennifer del Rosario-Malonzo. IBON International. https://iboninternational.org/download/state-terror-and-tyranny-in-the-philippines/?wpdmdl=3069&refresh=682d7e2e072281747811886.

Bolukbasi, Tolga, Kai-Wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. “Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings.” *Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS 2016)*. Conference on Neural Information Processing Systems, Barcelona, Spain. https://proceedings.neurips.cc/paper_files/paper/2016/file/a486cd07e4ac3d270571622f4f316ec5-Paper.pdf.

Bradshaw, Samantha, and Philip N. Howard. 2017. “Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation.” *Computational Propaganda Research Project*, Working Paper No. 2017.12, University of Oxford Internet Institute. https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/07/Troops-Trolls-and-Troublemakers.pdf.

Brown, Deborah. 2020. *Big Tech’s Heavy Hand Around the Globe*. Human Rights Watch, 8 September 2020. https://www.hrw.org/news/2020/09/08/big-techs-heavy-hand-around-globe.

Caballar, Rina Diane. 2024. “10 AI Dangers and Risks and How to Manage Them.” IBM, 3 September 2024. https://www.ibm.com/blog/10-ai-dangers-and-risks-and-how-to-manage-them/.

Caballero-Anthony, Mely, and Ralf Emmers, R. 2006. “Understanding the Dynamics of Securitizing Non-Traditional Security.” In *Non-Traditional Security in Asia: Dilemmas in Securitization*, edited by Mely Caballero-Anthony and Ralf Emmers. Ashgate.

Cabañes, Jason Vincent A., and Jayeel S. Cornelio. 2017. “The Rise of Trolls in the Philippines (And What We Can Do About It).” In *A Duterte Reader: Critical Essays on Rodrigo Duterte’s Early Presidency*, edited by Nicole Curato. Ateneo de Manila University Press.

Cadwalladr, Carole, and Emma Graham-Harrison. 2018. “How Cambridge Analytica Turned Facebook “likes” into a Lucrative Political Tool.” *The Guardian*, 17 March 2018. https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm#:~:text=By%20early%202014%2C%20Cambridge%20Analytica,university%2C%20but%20echoing%20Kosinski's%20work.

Companies Market Cap. 2025. “Large Tech Companies by Market Cap.” CompaniesMarketcap.Com, 28 May 2025. https://companiesmarketcap.com/tech/largest-tech-companies-by-market-cap/

Concepcion, Ronnie S., Rhen Anjerome R. Bedruz; Alvin B. Culaba; Elmer P. Dadios; Athena Rosz Ann R. Pascua. 2019. “The Technology Adoption and Governance of Artificial Intelligence in the Philippines.” *Proceedings of the 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control,*

Environment, and Management (HNICEM). 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Laoag, Philippines. https://doi.org/10.1109/HNICEM48295.2019.9072725.

Crawford, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Cueto, Santiago, María Balarin, Mauricio Saavedra, and Claudia Sugimaru. 2023. *Ed-tech in the Global South: Research Gaps and Opportunities*. Occasional Paper Series no. 91. Southern Voice. https://southernvoice.org/wp-content/uploads/2023/11/Ed-tech-Global-South-Cueto-et-al.-2023.pdf.

Curato, Nicole. 2021. "Social Manipulation and Disinformation in the Philippines." In *The Rise of Digital Repression*, edited by Steven Feldstein. Oxford University Press.

Dalofin, Raiza. 2022. "Introducing Issues Surrounding Public Education in the Philippines and Future Development." CPI. https://www.kofice.or.kr/hellocpi/html/05_issue/issue_12_01.html.

Davies, Pascale. 2024. "EU AI Act Reaction: Tech Experts Say the World's First AI Law Is 'Historic' but 'Bittersweet.'" *Euronews*, 16 March 2024. https://www.euronews.com/next/2024/03/16/eu-ai-act-reaction-tech-experts-say-the-worlds-first-ai-law-is-historic-but-bittersweet.

Divina, Nilo T., and Jay-R C. Ipac. 2024. "AI and the Law in the Philippines." *Asia Business Law Journal*, 15 April 2024. https://law.asia/ai-law-philippines/.

DOST-PCIEERD (Department of Science and Technology–Philippine Council for Industry, Energy and Emerging Technology Research and Development). 2021. *DOST PCIEERD ETDD: Artificial Intelligence and Information & Communications Technology: Roadmapping Executive Report*. DOST-PCIEERD. https://pcieerd.dost.gov.ph/images/pdf/2021/roadmaps/sectoral_roadmaps_division/etdd/Draft-1_AI--ICT-Roadmap-as-24.3.2021.pdf

DTI (Department of Trade and Industry). 2024. "DTI Drives AI Innovation with National Roadmap 2.0 and CAIR Launch." DTI, 5 July 2024. https://www.dti.gov.ph/archives/news-archives/dti-drives-ai-innovation-national-roadmap-2-0-cair-launch/.

The Economist. 2017. "The world's most valuable resource is no longer oil, but data." *The Economist*, 6 May 2017. https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.

Estrellado, Carie Justine P., and John Christian Miranda. 2023. "Artificial Intelligence in the Philippine Educational Context: Circumspection and Future Inquiries." *International Journal of Scientific and Research Publications* 13 (5): 16–22. https://doi.org/10.29322/IJSRP.13.05.2023.p13704.

Fairclough, Norman. 1992. *Discourse and Social Change*. Polity Press.

Fairclough, Norman. 1995. *Critical Discourse Analysis: The Critical Study of Language*. Longman Group Limited.

Fajardo, Margarita Felipe. 2016. "Teaching Critical Literacy Using Multimodal Texts to College Students in the Philippines." PhD diss., University of Wollongong. https://ro.uow.edu.au/articles/thesis/Teaching_Critical_Literacy_Using_Multimodal_Texts_to_College_Students_in_the_Philippines/27660906?file=50368947.

Foucault, Michel. 1972. *The Archaeology of Knowledge*. Pantheon Books.

Goh, Gillian. 2003. The "ASEAN Way": Non-Intervention and ASEAN's Role in Conflict Management. *Stanford Journal of East Asian Affairs* 3 (1): 113–18. https://www.academia.edu/3988485/113Gillian_Goh_Stanford_Journal_of_East_Asian_Affairs_GreaterEastAsia_The_ASEAN_Way_Non_Intervention_and_ASEAN_s_Role_in_Conflict_Managementin_conflicts_in_Haiti_and_Nicaragua_and_between.

Graham, Mark. 2011. "Time Machines and Virtual Portals: The Spatialities of the Digital Divide." *Progress in Development Studies* 11 (3): 211–27. https://doi.org/10.1177/146499341001100303.

Gutierrez, Kurt Lenard T., and Paulene Miriel L. Viacrusis. 2023. "Bridging the Gap or Widening the Divide: A Call for Capacity- Building in Artificial Intelligence for Healthcare in the Philippines." *Journal of Medicine University of Santo Tomas* 7 (2): 1325–34. https://doi.org/10.35460/2546-1621.2023-0081.

Han, Byung-Chul. 2018. *The Expulsion of the Other*. Polity Press.

Heaven, Will Douglas. 2023. "Geoffrey Hinton Tells Us Why He's Now Scared of the Tech He Helped Build." *MIT Technology Review*, 2 May 2023. https://www.technologyreview.com/2023/05/02/1072528/geoffrey-hinton-google-why-scared-ai/.

Kemp, Simon. 2022. "Digital 2022: The Philippines." Datareportal, 15 February 2022. https://datareportal.com/reports/digital-2022-philippines.

Kolkman, Daan. 2020. "'F**k the Algorithm'?: What the World Can Learn from the UK's A-Level Grading Fiasco." *LSE Blogs*, 26 August 2020. https://blogs.lse.ac.uk/impactofsocialsciences/2020/08/26/fk-the-algorithm-what-the-world-can-learn-from-the-uks-a-level-grading-fiasco/#:~:text=Last%20week%2C%20hundreds%20of%20students,exams%20due%20to%20the%20pandemic.

Kosinski, Michal, David Stillwell, and Thore Graepel. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Science of the United States of America* 110 (15): 5802–5. https://doi.org/10.1073/pnas.1218772110.

Kouadio, Louis, Ravinesh C. Deo, Vivekananda Byrareddy, Jan F. Adamowski, Shahbaz Mushtaq, and Van Phuong Nguyen. 2018. "Artificial Intelligence Approach for the Prediction of Robusta Coffee Yield Using Soil Fertility Properties." *Computers and Electronics in Agriculture* 155:324–448. https://doi.org/10.1016/j.compag.2018.10.014.

Kusaka, Wataru. 2017. *Moral Politics in the Philippines*. NUS Press & Kyoto University Press.

Kwet, Michael. 2017. "Operation Phakisa Education: Why a Secret? Mass Surveillance, Inequality, and Race in South Africa's Emerging National E-education System." *First Monday* 22 (12). https://doi.org/10.5210/fm.v22i12.8054.

Kwet, Michael. 2019. Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class* 60 (4): 3–26. https://doi.org/10.1177/0306396818823172.

Lee, Michael T., and Carol Theokary. 2021. "The Superstar Social Media influencer: Exploiting Linguistic Style and Emotional Contagion over Content?" *Journal of Business Research* 132:860–71. https://doi.org/10.1016/j.jbusres.2020.11.014.

McCoy, Alfred W. 2009. "Policing the Imperial Periphery: Philippine Pacification and the Rise of the U.S. National Security State." In *Colonial Crucible: Empire in the Making of the Modern American State,* edited by Alfred W. McCoy and Francisco A. Scarano. University of Wisconsin Press.

National Privacy Commission. 2016. Implementing Rules and Regulations of Republic Act No. 10173, Also Known as the "Data Privacy Act of 2012." Updated 12 July 2016. https://privacy.gov.ph/wp-content/uploads/2016/07/updated-draft-July-12-2016.pdf.

National Security Council. 2011. *2011–2016 National Security Policy: Securing the Gains of Democracy*. https://www.files.ethz.ch/isn/156790/Philippines%202011-2016.pdf

National Security Council. 2017. *2017-2022 National Security Policy*. https://nsc.gov.ph/attachments/article/NSP/NSP-2017-2022.pdf

National Security Council. 2023. *National Security Policy: 2023-2028*. https://nsc.gov.ph/images/NSS_NSP/National_Security_Policy_Manual_FINAL_E-COPY_with_WATERMARK_140823.pdf

Noujaim, Jehan, and KarimAmer, K., directors. 2019. *The Great Hack*. Netflix. https://www.netflix.com/ph-en/title/80117542.

OECD (Organisation for Economic Co-operation and Development. 2023. *PISA 2022 Results The State of Learning and Equity in Education*. Vol. 1. OECD Publishing. https://www.oecd.org/en/publications/pisa-2022-results-volume-i_53f23881-en.html.

Ordoñez, John Victor D. 2023. "Philippines Bats for Rules on Ethical Use of Autonomous Weapon Systems." *BusinessWorld,* 13 December 2023. https://www.bworldonline.com/the-nation/2023/12/13/563490/philippines-bats-for-rules-on-ethical-use-of-autonomous-weapon-systems/

Padilla, Len. 2014. "Four Ways the NSA Revelations Are Changing Businesses." *The Guardian,* 9 June 2014. https://www.theguardian.com/media-network/media-network-blog/2014/jun/09/edward-snowden-nsa-changing-business.

Rahman, Zara. 2023. *Machine Readable Me: The Hidden Ways Tech Shapes Our Identities*. 404 Ink Limited.

Ressa, Maria A. 2019. "Facebook Let My Government Target Me. Here's Why I Still Work with Them." *Time,* 17 January 2019. https://time.com/5505458/facebook-maria-ressa-philippines/.

Rosales, Marife A., Jo-ann V. Magsumbol; Maria Gemel B. Palconit; Alvin B. Culaba; Elmer P. Dadios. 2020. "Artificial Intelligence: The Technology Adoption and Impact

in the Philippines." *Proceedings of the 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*. 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), online. https://doi.org/10.1109/HNICEM51456.2020.9400025.

Roth, Emma. 2022. "Facebook's Plan to Offer Free Internet in Developing Countries Ended Up Costing Users, WSJ Reports." *The Verge*, 25 January 2022. https://www.theverge.com/2022/1/25/22900924/facebooks-free-internet-less-developed-costing-users-wsj.

Second Congressional Commission on Education (EDCOM II) Act. 2022. RA No. 11899.https://lawphil.net/statutes/repacts/ra2022/ra_11899_2022.html.

Sharman, Jon. 2018. "Metropolitan Police's Facial Recognition Technology 98% Inaccurate, Figures Show." *Independent*, 13 May 2018. https://www.independent.co.uk/news/uk/home-news/met-police-facial-recognition-success-south-wales-trial-home-office-false-positive-a8345036.html.

Shoemaker, Pamela J., and Gang (Kevin) Han. 2020. "The Supra-Gatekeepers: Gatekeeping in the Age of Social Media." *Communication & Society* 54:223–56. http://cschinese.com/word/32341062020.pdf.

Somers, James. 2021. "The Pastry A.I. That Learned to Fight Cancer." *New Yorker*, 18 March 2021. https://www.newyorker.com/tech/annals-of-technology/the-pastry-ai-that-learned-to-fight-cancer.

Stauffer, Brian. (2020). "Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Regaining Human Control." Human Rights Watch, 10 August 2020. https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and.

Strubell, Emma, Ananya Ganesh, and Andrew McCallum. 2019. "Energy and Policy Considerations for Deep Learning in NLP." *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy. Association for Computational Linguistics. https://doi.org/10.18653/v1/P19-1355.

Thompson, Mark R. 2010. "Reformism vs. Populism in the Philippines." *Journal of Democracy* 21 (4): 154–68. https://www.journalofdemocracy.org/articles/reformism-vs-populism-in-the-philippines/.

TikTok. 2024. "About Project Texas." TikTok U.S. Data Security. https://usds.tiktok.com/usds-about/.

Tobin, Meaghan. 2021. "How Facebook Discover Replicated Many of Free Basics' Mistakes." *Rest of World*, 8 June 2021. https://restofworld.org/2021/facebook-connectivity-discover/.

Tompa, Rachel. 2024. "Generative AI Develops Potential New Drugs for Antibiotic-Resistant Bacteria." Stanford Medicine News Center, 28 March 2024. https://med.stanford.edu/news/all-news/2024/03/ai-drug-development.html#:~:text=Generative%20AI%20

develops%20potential%20new%20drugs%20for%20antibiotic%2Dresistant%20bacteria,-share&text=Stanford%20Medicine%20researchers%20devise%20a,the%20drugs%20in%20the%20lab.

UNICEF (United Nations Children's Fund) and SEAMEO (Southeast Asian Ministers of Education Organization). 2020. *SEA-PLM 2019 Main Regional Report: Children's learning in 6 Southeast Asian Countries*. https://www.seaplm.org/PUBLICATIONS/regional%20results/SEA-PLM%202019%20Main%20Regional%20Report.pdf.

Wallerstein, Immanuel. 1976. "A World-System Perspective on the Social Sciences." *British Journal of Sociology* 27 (3): 343–52. https://doi.org/10.2307/589620.

Wodak, Ruth, and Michael Meyer. 2016. "Critical Discourse Studies: History, Agenda, Theory and Methodology." In *Methods of Critical Discourse Studies*, edited by Ruth Wodak and Michael Meyer. Sage.

Yu, Danni, Hannah Rosenfeld, and Abishek Gupta. 2023. "The 'AI Divide' Between the Global North and the Global South." World Economic Forum, 16 January 2023. https://www.weforum.org/stories/2023/01/davos23-ai-divide-global-north-global-south/.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.

THE UP CIDS DISCUSSION PAPER SERIES

The UP CIDS Discussion Paper Series features preliminary researches that may be subject to further revisions and is circulated to elicit comments and suggestions for enrichment and refinement. They contain findings on issues that are aligned with the core agenda of the research programs under the University of the Philippines Center for Integrative and Development Studies (UP CIDS).

CENTER FOR INTEGRATIVE AND DEVELOPMENT STUDIES

Established in 1985 by University of the Philippines (UP) President Edgardo J. Angara, the UP Center for Integrative and Development Studies (UP CIDS) is the policy research unit of the University that connects disciplines and scholars across the several units of the UP System. It is mandated to encourage collaborative and rigorous research addressing issues of national significance by supporting scholars and securing funding, enabling them to produce outputs and recommendations for public policy.

The UP CIDS currently has twelve research programs that are clustered under the areas of education and capacity building, development, and social, political, and cultural studies. It publishes policy briefs, monographs, webinar/conference/forum proceedings, and the Philippine Journal for Public Policy, all of which can be downloaded free from the UP CIDS website.

THE PROGRAM

The **Decolonial Studies Program (DSP)** focuses on the varying dimensions of coloniality/modernity that continue to impact the Global South societies and hinder their institutions from achieving their liberating potential. The term modernity/coloniality is often used with reference to continuing and often deleterious processes, conditions, and altitudes brought about by the colonial period that, Anibal Quijano and Walter Mignolo argue, are inextricably linked to the epistemology of modernity.

EDITORIAL BOARD

PROGRAM EDITORS

EDITORIAL STAFF

Get your policy papers published. Download open-access articles.

The Philippine Journal of Public Policy: Interdisciplinary Development Perspectives (PJPP), the annual peer-reviewed journal of the UP Center for Integrative and Development Studies (UP CIDS), welcomes submissions in the form of full-length policy-oriented manuscripts, book reviews, essays, and commentaries. The PJPP provides a multidisciplinary forum for examining contemporary social, cultural, economic, and political issues in the Philippines and elsewh ere. Submissions are welcome year-around.

For more information, visit cids.up.edu.ph. All issues/articles of the PJPP can be downloaded for free.

Get news and the latest publications.

Join our mailing list: bit.ly/signup_cids to get our publications delivered straight to your inbox! Also, you'll receive news of upcoming webinars and other updates.

We need your feedback.

Have our publications been useful? Tell us what you think: bit.ly/dearcids.

UNIVERSITY OF THE PHILIPPINES
CENTER FOR INTEGRATIVE AND DEVELOPMENT STUDIES
Lower Ground Floor, Ang Bahay ng Alumni, Magsaysay Avenue
University of the Philippines Diliman, Quezon City 1101

Telephone (02) 8981-8500 loc. 4266 to 4268
(02) 8426-0955

Email cids@up.edu.ph
cidspublications@up.edu.ph

Website cids.up.edu.ph