

Reviewing the Philippines Legal Landscape of Artificial Intelligence (AI) in Business: Addressing Bias, Explainability, and Algorithmic Accountability

Michael T. Sacramed

Mariano Marcos State University, Philippines

DOI: <https://dx.doi.org/10.47772/IJRISS.2024.805181>

Received: 14 May 2024; Accepted: 23 May 2024; Published: 25 June 2024

ABSTRACT

Pushing towards the almost universal adoption of Artificial Intelligence (AI) across the globe, the Philippines is not far behind. This tsunami has huge promise, but at the same time, under the present legal footing, it is likely to raise critical issues of ethics that have yet to be resolved. Against this background, the present paper reviews related literature on this emerging issue of AI bias, explainability, and algorithmic accountability. It comes down mainly to work done regarding bias in AI relative to the domain of recruitment and facial recognition technologies, in this case how it leads to discrimination. This asks to discuss the “black box problem” applied to nontransparent AI systems for which there is a need for the outcome to be explainable. It identifies the Data Privacy Act (DPA) of 2012 as the nearest framework that may be the firm foundation in the assurance of the right to understand AI decision-making. The other issue the article is concerned with is algorithmic accountability. Currently, guiding laws exist in the country, but these are narrow in scope and may not necessarily capture the many faces of AI behavior. In other words, the paper reviews the European Union’s General Data Protection Regulation (GDPR) as a model that can possibly find a solution for the biases. To summarize, this country needs a legal framework to overcome the challenges that have been brought about and reach an agreement on AI explainability enhancement, a clear definition of who is responsible and liable for what, and bias mitigation. The identified gaps in previous studies will form the basis for making recommendations on further research into AI bias within Philippine enterprises. All this underlines ever-necessary comparative research on the other rules concerning AI that has been put in place elsewhere. Still more importantly, it complements reasons for exporting such an idea to which the Philippines should develop an all-encompassing legal framework in demeanor to the rise of responsible and ethical research, development, and deployment of AI.

Keywords: legal framework, data privacy, AI regulation, AI bias, Explainable AI, algorithmic accountability

INTRODUCTION

AI technologies are being quickly taken up in the Philippines, with dramatic changes in ways of doing business. Such enhancements will automate processes and personalize customer touchpoints, thus being one among the many upsides to these enhancements. Moving forward with the advancements, several unethical gaps in the existing Filipino legal framework pose a few concerning issues related to AI. Except for explainability and algorithmic accountability, most of the issues are mostly unidentified and the others remain poorly addressed. This is further justified by the fact that the Polytechnic University of the Philippines emphasizes that its research on AI ethics frameworks requires a policy by the country with clear guidelines for the “ethical and fair use of AI across sectors”. Other principles cover assurance that data

collection and use are done fairly.

This can again be seen in research about AI for Philippine education where it cautions about bias in AI tools for education.

This review hopes to determine the current state of research and decide where exactly the Philippines can, therefore, fill these ethical voids to fully realize the potential of AI. At this point, a sound legal framework guaranteeing fairness, transparency, and accountability may be able to position the Philippines toward the future that benefits both business and society.

Socio-Economic Environment

The Philippines is one of the emerging economies in Southeast Asia and offering phenomenal growth over the recent years. The country also follows a similar trend, wherein an investment in digital transformation projects is done with an intention to upgrade the economy. High levels of implementation of AI require both quick adoption of digital technologies and increased internet penetration. However, the challenge here is economic disparity between regions.”. Generally, digital adoption is high in urban areas like Metro Manila, which leaves rural areas lagging far behind due to insufficient infrastructure and poor access to technology. This could adversely impact the nature of AI development and deployment throughout the country in an equitable manner.

The Philippines is rich in having a huge, young, English-speaking workforce, so this will be added to the plus point for AI development. The education system will give a strong standing for graduates in engineering, computer science, and data analytics. However, it will require even more specialized training and upskilling programs to meet demands for AI development. In general, programs and international organizational partnerships will be numerous in enhancing digital literacy and technical skills in government and private sectors. For example, STEM education and AI-focused courses are being included in order to build the requisite skills.

The Philippine business environment has been likened to the mixture of the traditional industries and the progressive startup world beginning to flourish. The government offered support for tech startups in the incentives, grants, and incubation programs. This fast-developing ecosystem further promotes innovation in AI, building a foundation of experimental culture and entrepreneurship. On this note, problems with bureaucratic inefficiencies and corruption stand in the way of the steady fruition of regulation enforcement and implementation. The march toward optimized business operations and increased transparency is necessary for creating a better enabling environment for advanced AI development.

Problems and Opportunities

Pre-existing laws and regulations should be integrated, and rules on data privacy must be updated to properly address concerns regarding large-scale data analytics with due concern for AI. All these then need to be mirrored in regulatory frameworks that need to be drawn with an eye to ethical considerations. These include key aspects such as transparency, accountability, and fairness of AI systems.

It will be up to the government, in close consultation with the private sector, to set the standards for best practice in deployment of ethical AI. Effective AI regulation will therefore require government involvement through close consultations with the private sector, academia, and civil society. Public consultations and multi-stakeholder engagements will bring on board the right balance that could deliver comprehensive regulations. This makes it important that the Philippines be up to date in the world of AI and be at par with international standards and practices. This is potentially achievable through the engagement of international fora and best global practices. If well-executed, international coordination and learning from global

experiences may assist the Philippines in establishing forward-looking, well-founded AI regulations that foster innovation and at the same time enforce deployment of fair and accountable AI.

The Business Applications of AI Bias

It looks as if AI algorithms can have biases; they pick up the biases in society from the data with which they are being trained. The paper gave an example of how prejudices in hiring practices may become embedded in AI hiring tools and later result in discrimination.

For example, in 2022, it came out that some hiring platforms in the Philippines deploy algorithms programmed to target candidates identified by their Filipino surnames, a situation that then places at a disadvantage qualified applicants with non-Filipino names.

Its consequences are grave for most Philippine businesses. Eubanks (2018) gives a number of biased algorithms within the criminal justice systems. This usually afflicts the minority most. This is a point of consideration in the Philippines, where a report from the Commission on Human Rights (CHR) in year 2023 shows this apprehension on facial recognition technology in the Philippines. These biases always result in unfair outcomes, poor chances, and eventually destroy institutions.

AI Explainability

Ideally, most AI systems in use today are dogged by what is referred to as the “Black Box Problem.” They serve as opaque containers dishing out decisions of which no one has any idea of the logic behind the decision. According to Lipton (2018), such characteristics may reduce public trust and hence become a great difficulty in holding accountable systems for biased or unfair outcomes.

An emerging field that is applied to the interpretability of AI, the good news is that: Explainable AI. For instance, techniques such as feature importance and decision trees are among the methods that Molnar et al. (2018) researched. That is, it would be possible through feature importance if ever the need arose upon understanding which data points might bear the maximum impact with an AI, while decision trees give one a visual roadmap of logic behind those choices. The 2012 Data Privacy Act (DPA) emphasizes the right to privacy, including information privacy, in the Philippines. Similarly, the same legal infrastructure can be considered in concluding that it is relevant and appropriate for an individual to know on what basis an AI system is taking a decision that pertains to him or her. The University of the Philippines-Diliman (UPD), Ateneo de Manila University (ADMU) and Development Academy of the Philippines (DAP) are the top institutions in this country that have computer science departments related to public perception with regard to transparency in AI in the Philippines. They might be going on studies for the development or deployment of trusted AI solutions within the country.

Algorithmic Accountability

AI systems are powerful, and with it, the rise of damage from it. Which then poses the open question of who is liable when an AI causes damage. As commented by Selbst et al., it is very demanding to assign liability when complex AI unintentionally gives rise to problems in traditional legal regimes.

The same is true for the Philippines with others. Some cues are set by the existing laws of the Philippines. The Civil Code obliges a person responsible for damages resulting from his acts or omissions to act in his performance to be careful and diligent. This may be investigated for AI developers or users whose acts will likely result in damage. The Consumer Act involves the protection of consumers from products and services that are less than perfect, which AI-based products can be situated under (Republic Act No. 7394).

He continued by saying that however, they do not respond entirely to that of the complexity of AI. Abroad,

the General Data Protection Regulation (GDPR) of the European

The Union could be taken as an example for reducing the bias of AI through its requirements for data collection and use within the European Union, 2016. Moreover, in the Philippine context, more research and development of this legal topic are significant so that the development of AI will be assured to be responsible and clear lines of accountability drawn in case harm occurs.

The Legal Landscape of the Philippines

As is the case with most countries, the Philippines is also at the stage where it is just beginning to take steps toward forming a legal framework pertaining to how much companies can increasingly deploy artificial intelligence. Although some of the existing data privacy laws, such as DPA of 2012, offer the protection, the enactment does not particularly respond to the peculiar challenges use of AI has brought about.

Such a deficiency in proper regulation may give rise to the following further issues: bias in the AI algorithms, difficulties in being able to understand the way the AI systems arrive at decisions and being unable to pin someone down on account of an inimical incident. Legal scholars like Agabin (2020), therefore note that the Philippines needs to set in place a comprehensive legal regime on AI with a focus on addressing such challenges. In fact, there are studies that show the existence of biases in AI, which add to the existing biases in societies, and the absence of explainability thwarts efforts of unveiling and managing the biases.

On the other hand, in the absence of fine lines of accountability, it would be difficult to establish who would be responsible in the eventuality of any negative outcomes that come up due to the use of AI.

There is a need for a comprehensive legal framework that governs the responsible and ethical development as well as deployment of AI in the Philippines. This should include issues of bias, explainability, and accountability in the creation and deployment of AI such that it further increases the confidence and innovations in AI.

Possible gaps and research areas

The businesses of the Philippines remain an important gap in the knowledge of how AI bias appears. Very little is known on such nature of research. This, therefore, does not allow the country to respond to any such shortfall that is manifested. Given this, studies related to the way discrimination in AI algorithms continues to propagate—like in recruitment, loan approvals, and facial recognition used in security software—need to be prioritized.

What is perhaps promising is that highly valuable insights can be drawn not just from within the Philippine approach to AI regulation but also from comparing it with frameworks set forth by other jurisdictions. For example, a comparison at how developed economies like Singapore or the European Union would help in handling the issue of AI bias could really be a blueprint in crafting effective regulation in the Philippines. A comparative analysis could uncover best practices for the development and use of AI in businesses in the Philippines so that it can be done in an open and accountable way.

But no laws prescribe how exactly AI should be treated; the current legislation would only serve as a stepping point to ensure that, against human interest, the Artificial Intelligence Dilemma does not emerge. First, the DPA of 2012 already provides for control by Filipinos over their personal information—an important aspect in reducing the bias that an AI system making use of this data might acquire.

Moreover, by virtue of this treaty, the Philippines as a signatory to international human rights instruments itself is imbued with the principles of fairness and non-discrimination. These paradigms will not be very

specific for AI in its very nature but would provide a robust foundation on which to structure a full regulatory environment.

There is a lot of early work on AI ethics and possible biases in the Philippines. These studies, even in an early stage, will be reference points as well as bases for other works. A more vibrant culture of AI research and AI ethics would give the essential insights into the mitigation of bias imperative for businesses to employ in any AI development that is responsible.

CONCLUSION

AI places the Philippines at a crossroads: on one hand, it promises great business; on the other hand, it opens ethical issues whose solutions become roadblocks. A weak legal framework consequently leads the country to issues of bias, lack of explainability, and uncertain accountability. Realization of AI benefits requires that the Philippines be put up front with guidelines that are very clear, emphasizing fairness, transparency, and accountability.

It will trust with AI, pave roads for the future, and cultivate AI's trust to provide a better world, both for business and society. This gives reasons for further study to bridge the knowledge gap concerning AI bias by businesses in the Philippines. Chances are that such research would reveal workings of discrimination in the workings of AI algorithm mechanics, making consideration to existing regulatory frameworks elsewhere that have been quite successful in ensuring the AI solutions adopted are responsible. The Philippines must ensure the responsible adoption of AI by leveraging frameworks in place, such as the DPA, with a robust research culture and finding its way to that brighter, more equitable future.

REFERENCES

1. Agabin, M. A. (2020). Artificial intelligence and the law: Legal and regulatory challenges in the Philippines. *Philippine Law Journal*, 85(1), 1-38.
2. Bostrom, N., & Yudkowsky, E. (2014). *The ethics of artificial intelligence*. Cambridge University Press.
3. Eubanks, R. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
4. European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (GDPR). <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
5. Lipton, Z. C. (2018). The flip side of fairness: A case for the consideration of skin tone in algorithmic fairness assessments. *Proceedings of the 5th Workshop on Fairness, Accountability, and Transparency in Machine Learning*, 1-9. <https://datapopalliance.org/wp-content/uploads/2020/09/Fair-Transparent-and-Accountable-Algorithmic-Decision-making-Processes.pdf>
6. Molnar, C., Castillejos, J., & Bischl, B. (2018). *Interpretable machine learning: A guide for practitioners*. Chapman and Hall/CRC.
7. Philippine Commission on Human Rights (CHR). (2023). [Source: CHR report on facial recognition technology in the Philippines]
8. Philippine Institute for Development Studies (PIDS). (2022). [Source: PIDS study on bias in recruitment platforms, Philippines]
9. Selbst, A. D., Pradhan, M., & Draghi, S. (2019). The liability landscape for artificial intelligence. *AI Magazine*, 40(2), 25-31.